Estatística de Redes Sociais

Antonio Galves

Módulo 3

11a. aula

Um modelo de rede com pressão social sobre os atores

Recapitulando: modelo de rede com pressão social

- Nestes modelos temos um **conjunto grande de atores** que interagem entre si. Todos os atores se influenciam mutuamente.
- Cada ator emite ao longo do tempo opiniões sobre um **assunto** específico. Suas opiniões podem ser ou *a favor* ou *contra* uma determinada posição.
- Cada vez que um ator emite uma opinião, ele zera suas convicções e fica à escuta das reações do grupo.
- Cada nova opinião que o ator emitir será influenciada pela opinião majoritária do grupo, manifestada depois dele ter emitido a sua última opinião. É assim que modelamos a pressão social sobre cada ator.

Recapitulando: modelo de rede com pressão social

- $A = \{1, 2, ..., N\}$: conjunto de atores.
- $\mathcal{O} = \{+1, -1\}$: conjunto de opiniões possíveis.
- $o_n \in \mathcal{O}$: n-ésima opinião omitida na rede.
- $A_n \in A$: ator que emitiu a n-ésima opinião.
- L_n^a: último instante em que o ator a emitiu opinião até o passo n.
- $L_n^a = \max(m \leq n : A_m = a)$.
- $U_n(a)$: pressão que a rede exerce sobre o ator a no instante n.
- $U_n(a) = \sum_{m=L_n^a+1}^{n} o_m$.
- $U_n = (U_n^1(1), \ldots, U_n^1(N))$: pressões exercidas pelo grupo social sobre todos os atores no instante n.

Recapitulando: modelo de rede com pressão social

- $$\mathbb{P}(o_{n+1} = o, A_{n+1} = a \mid U_n) = \frac{e^{oU_n(a)}}{\sum_{b \in A} \left[e^{(+1)U_n(b)} + e^{(-1)U_n(b)} \right]}.$$

- $$U_{n+1}(b) = \begin{cases} U_n(b) + o_{n+1} & \text{, se } b \neq A_{n+1}, \\ 0 & \text{, se } b = A_{n+1}. \end{cases}$$

Em linguagem de gente

- $$\mathbb{P}(o_{n+1} = o, A_{n+1} = a \mid U_n) = \frac{e^{oU_n(a)}}{\sum_{b \in A} \left[e^{(+1)U_n(b)} + e^{(-1)U_n(b)} \right]}.$$
- Essa probabilidade é proporcional a $e^{oU_n(a)}$.
- $e^{oU_n(a)}$ é grande se:
  - o e $U_n(a)$ forem ambos positivos ou ambos negativos;
  - e se o valor absoluto $|U_n(a)|$ for grande.
- Ou seja, se a pressão social sentida pelo ator a for grande num certo sentido, o mais provável é que sua próxima opinião vá no mesmo sentido.
- Observe que o valor absoluto $|U_n(a)|$ é grande se existir uma maioria de atores expressando opiniões no mesmo sentido, desde o instante L_n^a.

A opinião da rede

- Nesse modelo, a opinião o_{n+1} é fortemente influenciada pelos valores de U_n.
- Com efeito, qualquer que seja o ator que se expresse, a opinião emitida será o com probabilidade

$$\sum_{b \in A} \mathbb{P}(o_{n+1} = o, A_{n+1} = b \mid U_n) = \frac{\sum_{b \in A} e^{oU_n(b)}}{\sum_{b \in A} \left[e^{(+1)U_n(b)} + e^{(-1)U_n(b)} \right]}.$$

- O ator A_{n+1} que a expressa é escolhido com grande probabilidade dentre os atores b que tem $U_n(b)$ com o mesmo sentido de o e com grande valor absoluto de $U_n(b)$.

Características da cadeia $(U_n)_{n\geq 0}$

- $(U_n)_{n\geq 0}$ é uma cadeia de Markov com memória de alcance 1.
- Isso significa que para todo $n \geq 1$, a lista U_n depende só de U_{n-1} e do acaso.
- Em outras palavras, para sortearmos U_n utilizamos só os valores da lista U_{n-1}.
- Mais formalmente, para todo $n \geq 1$, e para toda escolha de listas com N componentes inteiros $z_0, z_1, ..., z_n$, temos

$$\mathbb{P}(U_n = z_n | U_{n-1} = z_{n-1}, \ldots, U_0 = z_0) = \mathbb{P}(U_n = z_n | U_{n-1} = z_{n-1}).$$

A propriedade de Markov em termos algorítmicos

- Em geral, a propriedade de Markov pode ser expressa em termos de um algoritmo para gerar a cadeia.
- O algoritmo tem a seguinte forma:
  1. Escolhemos o número de atores $N \geq 1$.
  2. Especificamos o conjunto $\mathcal{S}_N$ de listas que a cadeia pode assumir.
  3. Definimos uma função $f : \mathcal{S}_N \times [0,1] \to \mathcal{S}_N$.
  4. Inicialização: Escolhemos U_0 em $\mathcal{S}_N$ arbitrariamente.
  5. Para $n \geq 1$ fazemos

     $$U_n = f(U_{n-1}, \xi_n),$$

     onde $\xi_1, \xi_2, ...$ são variáveis aleatórias independentes entre si e uniformemente distribuídas no intervalo $[0,1]$.
- Em linguagem de gente, $\xi_1, \xi_2, ...$ são os números fornecidos por um gerador de números aleatórios.

Q U I Z

- Qual o problema com o algoritmo de geração descrito nesse pseudo-código?
  1. Escolhemos o número de atores $N \geq 1$.
  2. Especificamos o conjunto $\mathcal{S}_N$ de listas que a cadeia pode assumir.
  3. Definimos uma função $f : \mathcal{S}_N \times [0,1] \to \mathcal{S}_N$.
  4. Escolhemos U_0 em $\mathcal{S}_N$ arbitrariamente.
  5. Para $n \geq 1$ fazemos

     $$U_n = f(U_{n-1}, \xi_n),$$

     onde $\xi_1, \xi_2, \ldots$ são variáveis aleatórias independentes entre si e uniformemente distribuídas no intervalo $[0,1]$.

R E S P O S T A

- O problema é que esse pseudo-código rodará para sempre!
- Precisamos especificar até quando o processo de geração de listas vai continuar.

Exemplos

1. Escolhemos o número de atores $N \geq 1$ e escolhemos $T \geq 1$ o comprimento da amostra que queremos gerar.
2. Especificamos o conjunto $\mathcal{S}_N$ de listas que a cadeia pode assumir.
3. Definimos uma função $f : \mathcal{S}_N \times [0, 1] \to \mathcal{S}_N$.
4. Escolhemos U_0 em $\mathcal{S}_N$ arbitrariamente.
5. Para $n = 1, ..., T$ fazemos

$$U_n = f(U_{n-1}, \xi_n).$$

Exemplos

1. Escolhemos o número de atores $N \geq 1$.
2. Especificamos o conjunto $\mathcal{S}_N$ de listas que a cadeia pode assumir e especificamos um subconjunto de listas $\mathcal{R} \subset \mathcal{S}_N$.
3. Definimos uma função $f : \mathcal{S}_N \times [0, 1] \to \mathcal{S}_N$.
4. Escolhemos U_0 em $\mathcal{S}_N$ arbitrariamente.
5. Para $n \geq 1$ fazemos:
   5.1
   $$U_n = f(U_{n-1}, \xi_n).$$
   5.2 Se $U_n \in \mathcal{R}$, paramos o processo.

Conservar na memória, imprimir e fazer estatísticas

- Obviamente falta ainda uma coisa pratica. Conservar na memória as sucessivas listas $U_0, U_1, ...$ que estão sendo geradas.
- E fazer alguma coisa útil com elas: imprimi-las, fazer uma análise estatística da sequência, etc.

Q U I Z

- Lembramos as duas fórmulas centrais no modelo de rede social com pressão:
- $$\mathbb{P}(o_{n+1} = o, A_{n+1} = a \mid U_n) = \frac{e^{oU_n(a)}}{\sum_{b \in A} \left[e^{(+1)U_n(b)} + e^{(-1)U_n(b)} \right]}.$$
- $$U_{n+1}(b) = \begin{cases} U_n(b) + o_{n+1} & \text{, se } b \neq A_{n+1}, \\ 0 & \text{, se } b = A_{n+1}. \end{cases}$$
- Como podemos especificar o conjunto de listas $\mathcal{S}_N$ onde as listas $U_0, U_1, ...$ assumem seus valores?

RESPOSTA

- $\mathcal{S}_N \subseteq \mathbb{Z}^N$ com a seguinte restrição:
- $\mathcal{S}_N = \{(u(1), ..., u(N)) \in \mathbb{Z}^N : u(i) = 0,$ para algum $i = 1, ..., N\}$.

QUIZ

- O que acontece se, por inadvertência, escolhemos

$$U_0 = (U_0(1), ..., U_0(N)),$$

tal que $U_0(i) \neq 0$, para todo $i = 1, ..., N$?

R E S P O S T A

- Se $U_0 = (U_0(1), ..., U_0(N))$ for escolhido com todas as componentes diferentes de 0, a dinâmica do processo se encarrega de corrigir esse engano:
- $U_1 \in \mathcal{S}_N$, $U_2 \in \mathcal{S}_N$,...

As probabilidades de transição da cadeia $(U_n)_{n\geq 0}$

- Na teoria das cadeias de Markov, um papel importante é desempenhado pela matriz de probabilidades de transição da cadeia.
- Uma matriz de probabilidade de transição p para nosso modelo de rede social com pressão, é uma função

$$p : \mathcal{S}_N \times \mathcal{S}_N \to [0,1]$$

  assim definida:
- Para todo para $u, v \in \mathcal{S}_N$,

$$p(v|u) = \mathbb{P}(U_n = v | U_{n-1} = u).$$

Exercício de aquecimento

- Em um modelo com $N = 3$ atores, atribua valores para $p(v|u)$ onde u e v estão especificados abaixo.
  1. $u = (0, 0, 0)$ e $v = (0, +1, +1)$.
  2. $u = (0, 0, 0)$ e $v = (0, -1, -1)$.
  3. $u = (0, 0, 0)$ e $v = (0, +1, 0)$.
  4. $u = (1, -1, 0)$ e $v = (0, 0, -1)$.
- Que condições u e v tem que satisfazer para que $p(v|u) \neq 0$?
- Dados $z_0 = (0, 0, 0)$ e $z_3 = (+1, -1, 0)$, encontre z_1, z_2 em $\mathcal{S}_3$ tais que

$$\prod_{i=1}^{3} p(z_i|z_{i-1}) > 0.$$

Exercício básico sobre a cadeia $(U_n)_{n\geq 0}$

- É importante saber se uma cadeia de Markov é irredutível ou não.
- A definição de irredutibilidade é a seguinte:
- Se a cadeia assume valores no conjunto $\mathcal{S}$ e tem p como matriz de probabilidades de transição, diremos que a cadeia é irredutível se, para quaisquer dois elementos $u, v \in \mathcal{S}$ com $u \neq v$, existe $k \geq 1$ e uma sequência $z_0, ..., z_k$ de elementos de $\mathcal{S}$ tais que $z_0 = u$, $z_k = v$ e

$$p(z_{i+1}|z_i) > 0, \text{ para todo } i = 1, .., k.$$

- Diga se a cadeia $(U_n)_{n\geq 0}$ é irredutível em $\mathcal{S}_N$.

Exercício de preparação para o desafio

- Seja Y uma variável aleatória assumindo valores em $\mathcal{S} = \{1, 2, 3\}$ e tendo a seguinte distribuição de probabilidades:
- $\mathbb{P}(Y = 1) = \mathbb{P}(Y = 2) = 1/6$ e $\mathbb{P}(Y = 3) = 2/3$.
- Queremos simular a variável aleatória Y a partir de um número aleatório ξ distribuído uniformemente em $[0, 1]$, isto é, queremos que

  $$Y = f(\xi),$$

  onde $f : [0, 1] \to \mathcal{S}$.
- Como podemos construir essa função f?

Resposta

- Há muitas maneiras de definir $f : [0,1] \to \mathcal{S}$. Por exemplo

$$f(r) = \begin{cases} 1, \text{ se } 0 \leq r < 1/6, \\ 2, \text{ se } 1/6 \leq r < 2/6, \\ 3, \text{ se } 2/6 \leq r \leq 1. \end{cases}$$

- Assim, se definirmos $Y = f(\xi)$,

$$\mathbb{P}(Y = 1) = \mathbb{P}(0 \leq \xi < 1/6) = 1/6$$

$$\mathbb{P}(Y = 2) = \mathbb{P}(1/6 \leq \xi < 2/6) = 1/6$$

$$\mathbb{P}(Y = 3) = \mathbb{P}(2/6 \leq \xi \leq 1) = 2/3.$$

- No algoritmo para gerar a cadeia $(U_n)_{n\geq 0}$, um papel central é desempenhado pela função

$$f : \mathcal{S}_N \times [0,1] \to \mathcal{S}_N$$

que, a partir de uma lista inicial arbitrária U_0, gera sucessivamente as listas $U_1, U_2, ...$,

$$U_n = f(U_{n-1}, \xi_n),$$

sendo $\xi_1, \xi_2, ...$ variáveis aleatórias independentes entre si uniformemente distribuídas em $[0,1]$.

- D E S A F I O : Como podemos construir essa função f?

Exercícios

1. Seja Y uma variável aleatória assumindo valores em $\mathcal{S} = \{1, 2, 3\}$ e tendo a seguinte distribuição de probabilidades: $\mathbb{P}(Y = 1) = \mathbb{P}(Y = 2) = 1/6$ e $\mathbb{P}(Y = 3) = 2/3$. Encontre uma outra função $g : [0, 1] \to \mathcal{S}$, $g \neq f$, tal que seja possível gerar Y na forma $Y = f(\xi)$.
2. Seja p a matriz de probabilidade de transição para nosso modelo de rede social com pressão assumindo valores em $\mathcal{S}_N$:

   $$p : \mathcal{S}_N \times \mathcal{S}_N \to [0, 1];$$

   Para todo para $u, v \in \mathcal{S}_N$,

   $$p(v|u) = \mathbb{P}(U_n = v | U_{n-1} = u).$$

   Que condição $u, v \in \mathcal{S}_N$ tem que satisfazer para que $p(v|u) > 0$?

Exercícios

3. É importante saber se uma cadeia de Markov é irredutível ou não. A definição de irredutibilidade é a seguinte:
   Se a cadeia assume valores no conjunto $\mathcal{S}$ e tem p como matriz de probabilidades de transição, diremos que a cadeia irredutível para quaisquer dois elementos $u, v \in \mathcal{S}$ com $u \neq v$, existe $k \geq 1$ e um sequência $z_0, ..., z_k$ de elementos de $\mathcal{S}$ tais que $z_0 = u$, $z_k = v$ e
   $$p(z_{i+1}|z_i) > 0, \text{ para todo } i = 1, .., k.$$
   Diga se a cadeia $(U_n)_{n\geq 0}$ é irredutível em $\mathcal{S}_N$.