Estatística de Redes Sociais

Antonio Galves

Módulo 4

Polarização e evolução da rede com pressão social

19a. aula

Convergência para a medida de probabilidade invariante

Para que servem os Algoritmos 1 e 2

O Algoritmo 1 tem como objetivo colocar em prática os elementos utilizados na demonstração da Desigualdade 1.

1. Seja o conjunto de atores $\mathcal{A} = \{1, 2, ..., N\}$. Vamos primeiro ver que para qualquer $u \in \mathcal{S}_N$, a probabilidade

$$\mathbb{P}\left(|U_{n+N}(a)| \leq N-1, \text{ para todo } a \in \mathcal{A} \,\middle|\, U_n = u\right) \geq \left(\frac{1}{N}\right)^N.$$

O Algoritmo 2 tem como objetivo colocar em prática os elementos utilizados na demonstração da Desigualdade 2.

2. Depois vamos ver que existe $\epsilon_N \in (0, 1)$ tal que

$$\mathbb{P}\left(U_{n+N} = e^{(+1)} \,\middle|\, \bigcap_{a \in \mathcal{A}} \{|U_n(a)| \leq N-1\}\right) \geq \epsilon_N.$$

Definições usadas nos Algoritmos

- Dado $u \in \mathcal{S}_N$, para todo $a \in \{1, ..., N\}$ seja

$$r(a|u) = \frac{e^{+u(a)} + e^{-u(a)}}{\sum_{b \in A} \left[e^{(+1)u(b)} + e^{(-1)u(b)} \right]}.$$

- Dado $u \in \mathcal{S}_N$, seja

$$a_1(u) = \text{argmax}\{|u(a)| : a \in \mathcal{A}\},$$

$$a_2(u) = \text{argmax}\{|u(a)| : a \in \mathcal{A} \setminus \{a_1(u)\}\},$$

$$...$$

$$a_N(u) = \text{argmax}\{|u(a)| : a \in \mathcal{A} \setminus \{a_1(u), a_2(u), ..., a_{N-1}(u)\}\}.$$

Definições usadas nos Algoritmos

- Para todo $\xi \in [0,1]$, defina

$$A(u,\xi) = \inf\{a \in \mathcal{A} : \sum_{b=1}^{a} r(a_b(u)|u) > \xi\}.$$

- Dado $u \in \mathcal{S}_N$ e $a \in \{1, ..., N\}$, para todo $\xi \in \{0,1\}$ seja

$$O(u(a),\xi) = \begin{cases} +1, & \text{se } \xi < \dfrac{1}{1+e^{-2u(a)}}, \\ -1, & \text{em caso contrário.} \end{cases}$$

Algoritmo Preliminar

1. Inicialização: Escolha uma lista $u \in \mathcal{S}_N$. Escolha um tempo $T \geq 1$.
2. Defina $U_0 = u$ e $H_p = (u)$
3. Para n de 1 até T faça:
   3.1 Sorteie (independentemente dos sorteios anteriores) ξ_n^1 uniforme em $[0, 1]$
   3.2 Atualize $A_n = A(U_{n-1}, \xi_n^1)$.
   3.3 Sorteie (independentemente dos sorteios anteriores) ξ_n^2 uniforme em $[0, 1]$.
   3.4 Atualize $O_n = O(U_{n-1}(A_n), \xi_n^2)$.
   3.5 Defina $U_n = \pi_{A_n, O_n}(U_{n-1})$.
   3.6 Atualize H_p concatenando a sequência anterior H_p com U_n: $H_p = (H_p, U_n)$.
4. Imprima H_p.

Algoritmo 1

1. Inicialização: Escolha duas listas $u \in \mathcal{S}_N$ e $v \in \mathcal{S}_N$.
2. Defina $U_0^u = u, U_0^v = v$ e $H_1 = (U_0^u, U_0^v)$.
3. Para $n = 1, ..., N$, faça:
   3.1 Sorteie (independentemente dos sorteios anteriores) ξ_n^1 uniforme em $[0, 1]$
   3.2 Atualize $A_n^u = A(U_{n-1}^u, \xi_n^1)$ e $A_n^v = A(U_{n-1}^v, \xi_n^1)$.
   3.3 Sorteie (independentemente dos sorteios anteriores) ξ_n^2 uniforme em $[0, 1]$.
   3.4 Atualize $O_n^u = O(U_{n-1}^u(A_n^u), \xi_n^2)$ e $O_n^v = O(U_{n-1}^v(A_n^v), \xi_n^2)$.
   3.5 Defina $U_n^u = \pi_{A_n^u, O_n^u}(U_{n-1}^u)$ e $U_n^v = \pi_{A_n^v, O_n^v}(U_{n-1}^v)$.
   3.6 Atualize H_1 concatenando a sequência anterior H_1 com (U_n^u, U_n^v): $H_1 = (U_0^u, U_0^v) \ldots (U_n^u, U_n^v)$.
   3.7 Imprima H_1.

Definições usadas no Algoritmo 2

- Seja

$$l_1 = (N, N-1, ..., 1),$$

$$l_2 = (N-1, N-2, ..., 1, N),$$

$$...$$

$$l_N = (1, N, ..., 3, 2).$$

- Note que para $l_j(b)$ é o b-ésimo elemento da sequência l_j.
- Para todo $\xi \in [0,1]$ e $j \in \mathcal{A}_N$, defina

$$\alpha_j(u,\xi) = \inf\{a \in \mathcal{A} : \sum_{b=1}^{a} r(l_j(b)|u) > \xi\}.$$

Algoritmo 2

1. Inicialização: Escolha duas listas $u \in \mathcal{S}_N$ e $v \in \mathcal{S}_N$.
2. Defina $U_0^u = u, U_0^v = v$ e $H_2 = (U_0^u, U_0^v)$.
3. Para $n = 1, ..., N$, faça
   3.1 Sorteie (indep. dos sorteios anteriores) ξ_n^1 uniforme em $[0, 1]$
   3.2 Atualize $A_n^u = \alpha_n(U_{n-1}^u, \xi_n^1)$ e $A_n^v = \alpha_n(U_{n-1}^v, \xi_n^1)$.
   3.3 Sorteie (indep. dos sorteios anteriores) ξ_n^2 uniforme em $[0, 1]$.
   3.4 Atualize $O_n^u = O(U_{n-1}^u(A_n^u), \xi_n^2)$ e $O_n^v = O(U_{n-1}^v(A_n^v), \xi_n^2)$.
   3.5 Defina $U_n^u = \pi_{A_n^u, O_n^u}(U_{n-1}^u)$ e $U_n^v = \pi_{A_n^v, O_n^v}(U_{n-1}^v)$.
   3.6 Atualize H_2 concatenando a sequência anterior H_2 com (U_n^u, U_n^v):
   $H_2 = (U_0^u, U_0^v) \dots (U_n^u, U_n^v)$.
   3.7 Imprima H_2.

Concatenando os Algoritmos 1 e 2

- Iniciamos com duas listas $u, v \in \mathcal{S}_N$, $u \neq v$.
- Nos instantes $1, 2, ..., N$ utilizamos o Algoritmo 1.
- Nos instantes $N+1, N+2, ..., 2N$ utilizamos o Algoritmo 2, e assim sucessivamente.
- Formalmente, para $k \geq 0$,
- Nos instantes $2kN+1, 2kN+2, ..., 2kN+N$, utilizamos o Algoritmo 1.
- Nos instantes $(2k+1)N+1, (2k+1)N+2, ..., (2k+1)N+N$, utilizamos o Algoritmo 2.
- Se
$$U^u_{(2k+1)N+N} = U^v_{(2k+1)N+N} = e^{(+1)},$$
saímos do laço.
- Caso contrário, continuamos o laço.

D E S A F I O

- Escreva o pseudo-código do Algoritmo concatenando os Algoritmos 1 e 2 e terminando com o Algoritmo Preliminar, após as duas listas ficarem iguais à escada simultaneamente.

R E S P O S T A

1. Inicialização: Escolha duas listas $u \in \mathcal{S}_N$ e $v \in \mathcal{S}_N$.
2. Defina $U_0^u = u, U_0^v = v$, $T^e = \infty$, $H = (u, v)$.
3. Para $k = 0, ..., M$, faça:
   4.1 Se $U_{k2N}^u = U_{k2N}^v = e^{(+1)}$, atualize $T^e = 2kN$ e saia do laço.
   4.2 Se $k < M$:
      4.2.1 Use o Algoritmo 1 com listas iniciais (U_{2kN}^u, U_{2kN}^v) e faça $H = (H, H_1)$.
      4.2.2 Use o Algoritmo 2 com listas iniciais $(U_{2kN+N}^u, U_{2kN+N}^v)$ e faça $H = (H, H_2)$.
4. Se $k < M$, use o Algoritmo Preliminar com lista inicial $e^{(+1)}$ e número de iterações $(M - k)2N$ e faça $H = (H, (H_p, H_p))$, onde H_p é o histórico do algoritmo preliminar.
5. Se $T^e = \infty$, imprima "A concatenação dos Algoritmos 1 e 2, para $k = 0, ..., M$, não gerou $U_{k2N}^u = U_{k2N}^v = e^{(+1)}$ "
6. Imprima H.

Q U I Z

- Por que concatenamos os Algoritmos 1, 2 e Preliminar?

QUIZ

- Construímos simultaneamente $(U_n^u, U_n^v)_{n\geq 0}$ utilizando a concatenação dos Algoritmos 1, 2 e Preliminar que acabamos de descrever.
- Encontre um majorante para

$$\mathbb{P}(U_n^u \neq U_n^v).$$

RESPOSTA

- Vamos chamar de $T^{u,v}$ o primeiro instante n tal que $U_n^u = U_n^v$.
- Dizer que $U_n^u \neq U_n^v$ é equivalente a dizer que $T^{u,v} > n$.
- Ou seja,
$$\mathbb{P}(U_n^u \neq U_n^v) = \mathbb{P}(T^{u,v} > n).$$
- A concatenação dos Algoritmos 1 e 2 que acabamos de descrever, torna possível que $U_{2kN}^u = U_{2kN}^v = e^{(+1)}$, com $k = 0, 1, \dots$.
- Definimos $T^e = 2kN$, onde k é o primeiro índice de bloco de tamanho $2N$ no qual essa igualdade acontece.
- Obviamente, $T^{u,v} \leq T^e$.
- Ou seja,
$$P(T^{u,v} > n) \geq P(T^e > n).$$

QUIZ

- Construímos simultaneamente $(U_n^u, U_n^v)_{n\geq 0}$ utilizando a concatenação dos Algoritmos 1, 2 e Preliminar que acabamos de descrever.
- Encontre um majorante para

$$\mathbb{P}(T^e > 2kN).$$

RESPOSTA

- $T^e > 2kN$ significa que nas primeiras k iterações do laço do Algoritmo Concatenado não obtivemos o efeito almejado.
- Qual é o efeito almejado em cada iteração do laço?
- O efeito almejado é que o Algoritmo 1 faça em N passos a pressão da rede sobre todos os atores ficar, simultaneamente nas duas listas, entre $-N$ e N,
- e que em seguida o Algoritmo 2 faça com que nas duas redes, os atores $N, N-1, ..., 1$ emitam simultânea e sucessivamente a opinião $+1$.

- A probabilidade de que esses dois resultados sejam obtidos sucessivamente em um laço, é maior ou igual a

$$\left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1+e^{2(N-n)}},$$

onde

$$\zeta_n = \frac{2}{2+\sum_{k=0}^{n-1}[e^k+e^{-k}]+(N-n)(e^{(N-1)+n}+e^{-((N-1)+n)})}.$$

- Portanto, a probabilidade de que isso não ocorra em um laço, é majorada por

$$1-\left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1+e^{2(N-n)}}.$$

- A probabilidade de que o efeito almejado não ocorra em um laço, é majorada por

$$1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}.$$

- Como as variáveis $(\xi_1^1, \xi_1^2), (\xi_2^1, \xi_2^2), ..., (\xi_k^1, \xi_k^2)$ que determinam se o efeito almejado é alcançado são independentes entre si, os eventos "o efeito almejado não é alcançado nos laços 1,2,...,k" são todos independentes entre si.
- Logo,

$$\mathbb{P}(T^e > 2kN) \leq \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^k.$$

Q U I Z

- Construímos simultaneamente $(U_n^u, U_n^v)_{n\geq 0}$ utilizando a concatenação dos Algoritmos 1, 2 e Preliminar que acabamos de descrever.
- Quanto vale

$$\lim_{k\to+\infty} \mathbb{P}(T^e > 2kN)?$$

- Quanto vale

$$\lim_{n\to+\infty} \mathbb{P}(T^{u,v} > n)?$$

R E S P O S T A

- Pelo QUIZ anterior,

$$\mathbb{P}(T^e > 2kN) \leq \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^k .$$

- Logo,

$$\lim_{k \to +\infty} \mathbb{P}(T^e > 2kN) \leq \lim_{k \to +\infty} \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^k = 0.$$

- Para a segunda resposta, seja $\lfloor \frac{n}{2N} \rfloor$ o maior inteiro menor ou igual a $\frac{n}{2N}$.
- Logo, $n \geq 2\lfloor \frac{n}{2N} \rfloor N$ e

$$\mathbb{P}(T^{u,v} > n) \leq \mathbb{P}(T^{u,v} > 2\lfloor n/2N \rfloor N).$$

- Já vimos que

$$\mathbb{P}(T^{u,v} > 2kN) \leq \mathbb{P}(T^{e} > 2kN).$$

- Logo,

$$\lim_{n \to +\infty} \mathbb{P}(T^{u,v} > n) \leq \lim_{n \to +\infty} \mathbb{P}(T^{e} > 2\lfloor n/2N \rfloor N) = 0,$$

já que $\lim_{n \to +\infty} \lfloor n/2N \rfloor = +\infty$.

Essa majoração não depende de u, v

É importante observar que esta majoração não depende dos valores das listas iniciais u, v.

Ou seja,

$$\sup\{\mathbb{P}(T^{u,v} > n) : u, v \in \mathcal{S}_N\} \leq \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^{\lfloor n/2N \rfloor} .$$

QUIZ

- Vamos supor que a lista inicial $U_0 = u$ seja sorteada em $\mathcal{S}_N$ com probabilidade $\nu(u)$, onde ν é uma medida de probabilidade em $\mathcal{S}_N$.
- Para todo $n \geq 1$, quanto vale a distribuição de U_n?
- Em particular, quanto vale a distribuição de U_n quando a lista inicial é escolhida usando μ, a medida de probabilidade invariante da cadeia?

RESPOSTA

- Para todo $z \in \mathcal{S}_N$,

$$\mathbb{P}(U_n = z) = \sum_{u \in \mathcal{S}_N} \mathbb{P}(U_n = z | U_0 = u)\mathbb{P}(U_0 = u) =$$

$$\sum_{u \in \mathcal{S}_N} \mathbb{P}(U_n = z | U_0 = u)\nu(u).$$

- Se $\nu = \mu$ que é a medida de probabilidade invariante da cadeia, então, por definição,

$$\mathbb{P}(U_n = z) = \mu(z).$$

Notação

- Vamos denotar por
$$(U_n^{\nu})_{n\geq 0}$$
a cadeia cuja lista inicial U_0 é escolhida usando a distribuição ν.

QUIZ

- Seja μ a distribuição invariante da cadeia, e seja v uma lista qualquer do conjunto $\mathcal{S}_N$.
- Construímos simultaneamente $(U_n^\mu, U_n^v)_{n\geq 0}$, utilizando a concatenação dos Algoritmos 1, 2 e Preliminar que acabamos de descrever.
- Se no instante n, $U_n^\mu = U_n^v$, qual é a distribuição de U_n^v?
- Em outras palavras, se no instante n, $U_n^\mu = U_n^v$, quanto vale $\mathbb{P}(U_n^v = z)$ para todo $z \in \mathcal{S}_N$?

RESPOSTA

- Por hipótese, $U_n^v = U_n^\mu$.
- Já vimos que $\mathbb{P}(U_n^\mu = z) = \mu(z)$.
- Portanto,

$$\mathbb{P}(U_n^v = z) = \mathbb{P}(U_n^\mu = z) = \mu(z).$$

Convergência em distribuição da rede com pressão

- Vamos verificar que

$$\lim_{n \to \infty} \mathbb{P}(U_n^v = u) = \mu(u),$$

  para todo $u, v \in \mathcal{S}_N$,
- sendo U_n^v a lista após n etapas da rede com pressão social, tendo v como lista inicial,
- e μ é a medida de probabilidade invariante da cadeia.
- Começamos observando que

$$|\mathbb{P}(U_n^v = u) - \mu(u)| = |\mathbb{P}(U_n^v = u) - \mathbb{P}(U_n^\mu = u)| =$$

$$|\sum_{z \in \mathcal{S}_N} \mathbb{P}(U_n^v = u, U_n^\mu = z) - \sum_{z \in \mathcal{S}_N} \mathbb{P}(U_n^v = z, U_n^\mu = u)|,$$

  onde $(U_n^u, U_n^\mu)_{n \geq 0}$ são construídos simultaneamente usando o Algoritmo concatenado.

QUIZ

- Que termos permanecem, quando calculamos a diferença,

$$|\sum_{z \in \mathcal{S}_N} \mathbb{P}(U_n^v = u, U_n^\mu = z) - \sum_{z \in \mathcal{S}_N} \mathbb{P}(U_n^v = z, U_n^\mu = u)|?$$

RESPOSTA

- $$\left|\sum_{z\in\mathcal{S}_N}\mathbb{P}(U_n^v=u,U_n^\mu=z)-\sum_{z\in\mathcal{S}_N}\mathbb{P}(U_n^v=z,U_n^\mu=u)\right|=$$

  $$\left|\sum_{z\neq u}\mathbb{P}(U_n^v=u,U_n^\mu=z)-\sum_{z\neq u}\mathbb{P}(U_n^v=z,U_n^\mu=u)\right|.$$

QUIZ

- Verifique que

$$|\sum_{z\neq u}\mathbb{P}(U_n^v=u,U_n^\mu=z)-\sum_{z\neq u}\mathbb{P}(U_n^v=z,U_n^\mu=u)|\leq\mathbb{P}(U_n^v\neq U_n^\mu).$$

R E S P O S T A

- $$|\sum_{z\neq u}\mathbb{P}(U_n^v = u, U_n^\mu = z) - \sum_{z\neq u}\mathbb{P}(U_n^v = z, U_n^\mu = u)| \leq$$

$$\sum_{z\neq u}\mathbb{P}(U_n^v = u, U_n^\mu = z) + \sum_{z\neq u}\mathbb{P}(U_n^v = z, U_n^\mu = u) =$$

$$\mathbb{P}(U_n^v = u, U_n^\mu \neq u) + \mathbb{P}(U_n^v \neq u, U_n^\mu = u) \leq$$

$$\sum_{(z_1,z_2),z_1\neq z_2} P(U_n^v = z_1, U_n^\mu = z_2) = \mathbb{P}(U_n^v \neq U_n^\mu).$$

Concluindo a verificação da convergência

- Utilizando os QUIZes anteriores, concluímos que

$$|\mathbb{P}(U_n^v = u) - \mu(u)| \leq \mathbb{P}(U_n^v \neq U_n^\mu) = \mathbb{P}(T^{v,\mu} > n),$$

$$\mathbb{P}(T^{v,\mu} > n) \leq \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^{\lfloor n/2N \rfloor}.$$

- Como

$$\lim_{n\to\infty} \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^{\lfloor n/2N \rfloor} = 0,$$

concluímos que

$$\boxed{\lim_{n\to+\infty} \mathbb{P}(U_n^v = u) = \mu(u).}$$

O que já sabemos

1. Para todo $u, v \in \mathcal{S}_N$, para todo $\beta \in [0, +\infty)$ vale a convergência

$$\lim_{n \to +\infty} \mathbb{P}(U_n^{\beta,v} = u) = \mu_\beta(u),$$

   que não depende da lista inicial v, só da lista final u.

   Nessa fórmula, μ_β é a medida invariante da cadeia $(U_n^{\beta,v})_{n \geq 0}$.

2. Sabemos também que $\mu_\beta(u) = \mu_\beta(-u)$, para todo $u \in \mathcal{S}_N$.
3. Usando 1 e 2, já vimos que

$$\lim_{n \to \infty} \mathbb{P}(O_n^{\beta,v} = +1) = 1/2,$$

   qualquer que seja a lista inicial v e para qualquer $\beta \in [0, \infty)$.

O que ainda não sabemos: o efeito de β sobre a rede

- Não sabemos se o valor do parâmetro de polarização β tem algum efeito sobre o comportamento da rede social ao longo do tempo.
- Será que a evolução da rede social tem características diferentes, conforme a polarização seja grande ou pequena?
- Essa é uma questão da maior importância, considerando a situação de grande polarização de muitas redes sociais atualmente.
- É essa a questão que abordaremos nas aulas finais do curso.

Exercícios

1. Descreva como evoluem $(U_n^{\beta,v})_{n\geq 0}, (A_n^{\beta,v})_{n\geq 0}, (O_n^{\beta,v})_{n\geq 0}$ no caso em que $\beta = 0$.
2. **Exercício dedicado aos matemáticos do grupo:** sabendo que

$$\sup\{\mathbb{P}(T^{u,v} > n) : u, v \in \mathcal{S}_N\} \leq \left[1 - \left(\frac{1}{N}\right)^N \prod_{n=1}^{N} \zeta_n \frac{1}{1 + e^{2(N-n)}}\right]^{\lfloor n/2N \rfloor} .$$

   encontre constantes $c, C > 0$ tais que

$$\sup\{\mathbb{P}(T^{u,v} > n) : u, v \in \mathcal{S}_N\}) \leq Ce^{-cn}.$$

3. **Exercício preparatório para as próximas aulas:** mostre que a equação

$$\frac{e^{\beta m} - e^{-\beta m}}{e^{\beta m} + e^{-\beta m}} - m = 0$$

   é satisfeita por um ou por três valores de m dependendo do valor de β. Sugestão: faça uma mudança de variáveis $\beta m = r$.