ε-regulární páry

Zdeněk Dvořák

29. listopadu 2016

Nechť A, B a C jsou disjunktní podmnožiny vrcholů nějakého grafu. Pak

$$\begin{aligned}
e(A,B) &= \text{počet hran s jedním koncem v } A \text{ a druhým v } B\\
d(A,B) &= \frac{e(A,B)}{|A||B|}\\
q(A,B) &= \text{počet 4-cyklů } v_1v_2v_3v_4, \text{ kde } v_1,v_3\in A \text{ a } v_2,v_4\in B\\
t(A,B,C) &= \text{počet trojúhelníků s jedním vrcholem v } A, \text{ druhým v } B \text{ a třetím v } C
\end{aligned}$$

1 Náhodné grafy

Nechť $0 \le p \le 1$. Nechť $G_{n,p}$ je náhodný graf na n vrcholech, obsahující každou hranu nezávisle s pravděpodobností p. Nechť A, B, C jsou disjunktní podmnožiny vrcholů $G_{n,p}$.

Pozorování 1.

$$\begin{aligned}
\mathbf{E}(e(A,B)) &= p|A||B|\\
\mathbf{E}(d(A,B)) &= p\\
\mathbf{E}(q(A,B)) &= p^4|A|(|A|-1)|B|(|B|-1)\approx p^4|A|^2|B|^2\\
\mathbf{E}(t(A,B,C)) &= p^3|A||B||C|.
\end{aligned}$$

Věta 2 (Černovova nerovnost)**.** *Nechť $0 \le p \le 1$. Nechť X_1, $\ldots$, X_n jsou nezávislé proměnné, rovné 1 s pravděpodobností p a rovné 0 s pravděpodobností $1-p$. Nechť $X=\sum_{i=1}^n X_i$. Pak pro každé $t\ge 0$,*

$$Pr(|X-pn|>t) < 2e^{-\frac{t^2}{2(pn+t/3)}}.$$

Lemma 3. *Nechť $\varepsilon>0$. S pravděpodobností $1-1/e^{\Omega(n^2)}$ platí, že všechny disjunktní podmnožiny A a B vrcholů $G_{n,p}$ velikosti alespoň εn splňují $|d(A,B)-p|\le\varepsilon$.*

Důkaz. Zafixujme nějaké takové A a B. Dle Věty 2,

$$\Pr(|e(A,B) - p|A||B|| > \varepsilon|A||B|) < 2e^{-\frac{\varepsilon^2}{2(p+\varepsilon/3)}|A||B|} \leq 2e^{-\frac{\varepsilon^4}{2(p+\varepsilon/3)}n^2}.$$

Položme $\kappa = \frac{\varepsilon^4}{2(p+\varepsilon/3)} > 0$, takže

$$\Pr(|d(A,B) - p| > \varepsilon) < 2e^{-\kappa n^2}.$$

Podmnožiny A a B lze zvolit méně než 4^n způsoby, takže pravděpodobnost, že nějaké dvě takové podmnožiny splňují $|d(A,B) - p| > \varepsilon$ je méně než

$$4^n 2e^{-\kappa n^2} = e^{(\log 4)n + \log 2 - \kappa n^2} = 1/e^{\Omega(n^2)}.$$

□

Definice 1. *Nechť G je graf a $\delta, \varepsilon > 0$. Pár (A,B) disjunktních podmnožin vrcholů G je (δ,ε)-*regulární*, jestliže*

- *všechny $A' \subseteq A$ a $B' \subseteq B$ takové, že $|A'| \geq \delta|A|$ a $|B'| \geq \delta|B|$, splňují*

$$|d(A',B') - d(A,B)| \leq \varepsilon.$$

*Pár je ε-*regulární*, je-li $(\varepsilon,\varepsilon)$-*regulární.

Lemma 4. *Nechť $\alpha, \varepsilon > 0$. Asymptoticky skoro jistě platí, že pro každé disjunktní podmnožiny A a B vrcholů $G_{n,p}$ velikosti alespoň αn je pár (A,B) ε-regulární.*

Důkaz. Dle Lemma 3 je pravděpodobnost, že A a B netvoří ε-regulární pár nejvýše $1/e^{\Omega(n^2)}$. Jelikož A a B lze zvolit pouze 4^n způsoby, pravděpodobnost, že nějaké dvě takové množiny netvoří ε-regulární pár je nejvýše $4^n/e^{\Omega(n^2)} = 1/e^{\Omega(n^2)}$. □

2 Vlastnosti ϵ-regulárních párů

Lemma 5. *Nechť G je graf a (A,B) je (δ,ε)-regulární pár v G pro nějaké $0 < \delta, \varepsilon \leq 1$. Pak*

- *počet vrcholů v A s více než $(d(A,B)+\varepsilon)|B|$ sousedy v B je méně než $\delta|A|$ a*
- *počet vrcholů v A s méně než $(d(A,B)-\varepsilon)|B|$ sousedy v B je méně než $\delta|A|$.*

Důkaz. Nechť A_1 je množina vrcholů v A s více než $(d(A,B)+\varepsilon)|B|$ sousedy v B. Pro spor předpokládejme, že $|A_1| \geq \delta|A|$. Jelikož (A,B) je (δ,ε)-regulární, platí $|d(A_1,B) - d(A,B)| \leq \varepsilon$. Ale na druhou stranu,

$$\begin{aligned} |d(A_1,B) - d(A,B)| &\geq \frac{e(A_1,B)}{|A_1||B|} - d(A,B) \\ &> \frac{|A_1|(d(A,B)+\varepsilon)|B|}{|A_1||B|} - d(A,B) = \varepsilon, \end{aligned}$$

což je spor. Druhá nerovnost se dokáže obdobně. □

Lemma 6. *Nechť G je graf a (A,B) je (δ,ε)-regulární pár v G pro nějaké $0 < \delta, \varepsilon \leq 1$. Nechť $B' \subseteq B$ má velikost alespoň $\delta|B|$. Pak*

- *počet vrcholů v A s více než $(d(A,B)+\varepsilon)|B'|$ sousedy v B' je méně než $\delta|A|$ a*
- *počet vrcholů v A s méně než $(d(A,B)-\varepsilon)|B'|$ sousedy v B' je méně než $\delta|A|$.*

Lemma 7. *Nechť G je graf a (A,B) je (δ,ε)-regulární pár v G pro nějaké $0 < \delta, \varepsilon \leq 1$. Nechť $|A| = |B| = n$ a $p := d(A,B)$. Pak*

$$|q(A,B) - p^4 n^4| \leq (2\delta + 15\varepsilon)n^4 + 2n^3.$$

Důkaz. Ukažme, že $q(A,B) \leq p^4n^4 + (2\delta+15\varepsilon)n^4$. Druhá nerovnost se dokáže obdobně.

Pro $v \in A$ si jako B_v označme množinu sousedů v v B. Nechť $A_1 = \{v \in A : |B_v| - pn > \varepsilon n\}$. Dle Lemma 5 máme $|A_1| < \delta n$. Jako A_2 označme množinu vrcholů $v \in V(G)$ takových, že $|B_v| < \delta n$. Pro $v \in A \setminus (A_1 \cup A_2)$ si jako A_v označme množinu vrcholů $w \in A$, které mají více než $(p+\varepsilon)|B_v|$ sousedů v B_v; jelikož $v \notin A_2$, máme $|B_v| \geq \delta n$, a proto Lemma 6 implikuje $|A_v| < \delta n$.

Počet 4-cyklů $v_1v_2v_3v_4$ takových, že $v_1 \in A_1$, je nejvýše $\delta n \cdot n^3$. Počet 4-cyklů $v_1v_2v_3v_4$ takových, že $v_1 \in A_2$, je nejvýše $n(\delta n)^2 n \leq \delta n^4$. Počet 4-cyklů $v_1v_2v_3v_4$ takových, že $v_1 \in A \setminus (A_1 \cup A_2)$ a $v_3 \in A_{v_1}$, je nejvýše $n \cdot \delta n \cdot n^2$. Počet 4-cyklů $v_1v_2v_3v_4$ takových, že $v_1 \in A \setminus (A_1 \cup A_2)$ a $v_3 \in A \setminus A_{v_1}$, je nejvýše $n \cdot n \cdot \big((p+\varepsilon)|B_{v_1}|\big)^2 \leq n \cdot n \cdot \big((p+\varepsilon)^2 n\big)^2$. Celkem tedy dostáváme nejvýše

$$\left(2\delta + (p+\varepsilon)^4\right) n^4 \leq p^4n^4 + (2\delta + 15\varepsilon)n^4$$

čtyřcyklů. □

Lemma 8. *Nechť G je graf a (A,B), (B,C) a (A,C) jsou (δ,ε)-regulární páry v G pro nějaké $0 < \delta, \varepsilon \leq \frac{1}{2}$. Jestliže $|A| = |B| = |C| = n$, pak*

$$|t(A,B,C) - d(A,B)d(B,C)d(A,C)n^3| \leq (6\delta + 7\varepsilon)n^3.$$

Důkaz. Ukážeme, že $t(A,B,C) \geq d(A,B)d(B,C)d(A,C)n^3 - (6\delta + 7\varepsilon)n^3$. Druhá nerovnost se dokáže obdobně.

Tvrzení je triviální, jestliže $\min(d(A,B), d(B,C), d(A,C)) \leq 6\delta + 7\varepsilon$. Předpokládejme tedy $d(A,B), d(B,C), d(A,C) > 6\delta + 7\varepsilon$. Jako A_1 si označme množinu vrcholů $v \in A$, které mají méně než $(d(A,B) - \varepsilon)n$ sousedů v B. Jako A_2 si označme množinu vrcholů $v \in A$, které mají méně než $(d(A,C) - \varepsilon)n$ sousedů v C. Dle Lemma 5 platí $|A_1| \leq \delta n$ a $|A_2| \leq \delta n$.

Nechť $v \in A \setminus (A_1 \cup A_2)$ a označme jako B_v a C_v množiny sousedů v v B a C. Máme $|B_v| \geq (d(A,B) - \varepsilon)n \geq \delta n$ a $|C_v| \geq (d(A,C) - \varepsilon)n \geq \delta n$, a jelikož (B,C) je (δ,ε)-regulární pár, máme $d(B_v, C_v) \geq d(B,C) - \varepsilon$.

Zvolíme-li libovolné $v \in A \setminus (A_1 \cup A_2)$ a libovolnou hranu mezi B_v a C_v, dostaneme trojúhelník. Tedy $t(A,B,C) \geq (1 - 2\delta)n \cdot (d(B,C) - \varepsilon)[(d(A,B) - \varepsilon)n][(d(A,C) - \varepsilon)n] \geq d(A,B)d(B,C)d(A,C)n^3 - (6\delta + 7\varepsilon)n^3$. □