

Bổ đề Szemerédi

Trương Phước Nhân, Cần Thơ, 03/08/2017

Cho G là một đồ thị, và $A, B \subset V(G)$. Ta định nghĩa mật độ của cặp (A, B) là tỉ số

$$d(A, B) := \frac{|E(A, B)|}{|A||B|} = \frac{|\{(x, y) \in A \times B : xy \in E(G)\}|}{|A \times B|}$$

Cặp (A, B) được gọi là ε -đều nếu với mọi $A' \subset A$ và $B' \subset B$ sao cho $|A'| \geq \varepsilon|A|$ và $|B'| \geq \varepsilon|B|$ thì ta luôn có $|d(A, B) - d(A', B')| \leq \varepsilon$.

Một phân hoạch $V(G) = X_1 \cup X_2 \cup \dots \cup X_m$ được gọi là ε -đều nếu

$$\sum \frac{|X_i||X_j|}{n^2} \leq \varepsilon, \text{ trong đó tổng được lấy trên các cặp } (X_i, X_j) \text{ không } \varepsilon\text{-đều}$$

Một cách tương đương thì đánh giá trên nói lên rằng phân hoạch $V(G) = X_1 \cup X_2 \cup \dots \cup X_m$ là ε -đều nếu xác suất để $(x, y) \in V(G)^2$ nằm trong một cặp (X_i, X_j) không ε -đều là nhỏ hơn hoặc bằng ε .

Cho trước một phân hoạch $V(G) = X_1 \cup X_2 \cup \dots \cup X_m$ ta định nghĩa trung bình bình phương mật độ bởi

$$\sum_{i,j=1}^m \frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2$$

Do $\sum_{i,j=1}^m \frac{|X_i||X_j|}{n^2} = 1$ và $0 \leq d(X_i, X_j) \leq 1$ nên giá trị của trung bình bình phương mật độ luôn nằm giữa 0 và 1

Kết quả 1: Cho $X_1, X_2, \dots, X_m$ là một phân hoạch của $V(G)$, và $Y_1, Y_2, \dots, Y_M$ là một phân hoạch mịn hơn của $X_1, X_2, \dots, X_m$ (hiểu theo nghĩa là với tập Y_i bất kì ta luôn tìm được một chỉ số j để $Y_i \subset X_j$). Khi đó, trung bình bình phương mật độ khi tính theo phân hoạch $Y_1, Y_2, \dots, Y_M$ sẽ luôn lớn hơn hoặc bằng với khi tính theo $X_1, X_2, \dots, X_m$.

Chứng minh:

Giả sử với mỗi chỉ số i , tập X_i được phân hoạch thành $X_{i_1} \cup X_{i_2} \cup \dots \cup X_{i_{t_i}}$. Khi đó:

$$\begin{aligned} d(X_i, X_j)^2 &= \left(\sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{|X_i||X_j|} d(X_{i_s}, X_{j_t}) \right)^2 \stackrel{\text{Cauchy-Schwartz}}{\leq} \left(\sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{|X_i||X_j|} \right) \left(\sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{|X_i||X_j|} d(X_{i_s}, X_{j_t})^2 \right) \\ &\Rightarrow d(X_i, X_j)^2 \leq \sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{|X_i||X_j|} d(X_{i_s}, X_{j_t})^2 \quad (\text{lưu ý: } \sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{|X_i||X_j|} = 1) \\ &\Rightarrow \frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2 \leq \sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{n^2} d(X_{i_s}, X_{j_t})^2 \\ &\Rightarrow \sum_{i,j} \frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2 \leq \sum_i \sum_{s,t} \frac{|X_{i_s}||X_{j_t}|}{n^2} d(X_{i_s}, X_{j_t})^2 \end{aligned}$$

Từ hệ thức trên ta suy ra điều phải chứng minh.

Nhận xét: Với cùng một phương pháp lập luận ta dễ dàng chỉ ra kết quả dưới đây:

Kết quả 2: Cho $G = (X \cup Y, E)$ là một đồ thị lưỡng phân và xét các phân hoạch $X = X_1 \cup X_2 \cup \dots \cup X_r$, $Y = Y_1 \cup Y_2 \cup \dots \cup Y_s$. Giả sử $Z_1 \cup Z_2 \cup \dots \cup Z_R$ và $W_1 \cup W_2 \cup \dots \cup W_S$ lần lượt là các phân hoạch mịn hơn của $X_1 \cup X_2 \cup \dots \cup X_r$ và $Y_1 \cup Y_2 \cup \dots \cup Y_s$. Khi đó:
$$\sum_{i=1}^r \sum_{j=1}^s \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 \leq \sum_{i=1}^R \sum_{j=1}^S \frac{|Z_i||W_j|}{|X||Y|} d(Z_i, W_j)^2$$

Kết quả 3: Giả sử $(X, Y) \subset V(G)^2$ sao cho (X, Y) không là ε -đều. Đặt $d(X, Y) = \alpha$. Khi đó, ta có thể tìm được các phân hoạch $X = X_1 \cup X_2$ và $Y = Y_1 \cup Y_2$ sao cho $\sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 \geq \varepsilon^4 + \alpha^2$

Chứng minh: Do (X, Y) không là ε -đều nên ta có thể tìm được $X_1 \subset X$ và $Y_1 \subset Y$ sao cho

$$|d(X_1, Y_2) - \alpha| \leq \varepsilon \text{ và } |X_1| \geq \varepsilon |X|, |Y_1| \geq \varepsilon |Y|.$$

Đặt $u(X_i, Y_i) = d(X_i, Y_i) - \alpha$. Khi đó:

$$\begin{aligned} \varepsilon^4 &\leq \frac{|X_1||Y_1|}{|X||Y|} u(X_1, Y_1)^2 \leq \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} u(X_i, Y_j)^2 \\ &= \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - 2\alpha \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j) + \alpha^2 \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} \\ &= \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - 2\alpha \sum_{i,j=1}^2 \frac{|E(X_i, Y_j)|}{|X||Y|} + \alpha^2 \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} \\ &= \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - 2\alpha \frac{|E(X, Y)|}{|X||Y|} + \alpha^2 \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} \\ &= \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - 2\alpha^2 + \alpha^2 \\ &\Rightarrow \varepsilon^4 + \alpha^2 \leq \sum_{i,j=1}^2 \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2. \end{aligned}$$

Từ đây ta suy ra điều phải chứng minh.

Kết quả 4:

Cho G là một đồ thị với n đỉnh, và gọi $X_1, X_2, \dots, X_m$ là một phân hoạch không ε -đều của $V(G)$. Khi đó ta có thể tìm được một phân hoạch mịn hơn $X_{11} \cup \dots \cup X_{1r_1} \cup X_{21} \cup \dots \cup X_{2r_2} \cup \dots \cup X_{m1} \cup \dots \cup X_{mr_m}$ sao cho $r_i \leq 2^{2^m}$ và trung bình bình phương mật độ khi tính theo phân hoạch này tăng ít nhất là ε^5 khi tính theo phân hoạch ban đầu.

Chứng minh:

Đặt $I = \{(i, j) : (X_i, X_j) \text{ không } \varepsilon\text{-đều}\}$. Giả sử α^2 là trung bình bình phương mật độ của đồ thị G tính theo phân hoạch $X_1, X_2, \dots, X_m$.

Với mỗi $(i, j) \in I$, từ kết quả 3) ta tìm được các phân hoạch $X_i = A_1^{ij} \cup A_2^{ij}$ và $X_j = B_1^{ij} \cup B_2^{ij}$ sao cho

$$\sum_{p,q=1}^2 \frac{|A_p^{ij}||B_q^{ij}|}{|X_i||X_j|} d(A_p^{ij}, B_q^{ij})^2 \geq d(X_i, X_j)^2 + \varepsilon^4$$

Bây giờ với mỗi chỉ số i , gọi $X_{i1} \cup X_{i2} \cup \dots \cup X_{ir_i}$ là phân hoạch của X_i thành tối đa 2^{2^m} tập sao cho mỗi A_1^{ij}, A_2^{ij} là hợp của một số X_{ih} . Sử dụng kết quả 2) ta thu được :

$$\sum_{p=1}^{r_i} \sum_{q=1}^{r_j} \frac{|X_{ip}||X_{jq}|}{|X_i||X_j|} d(X_{ip}, X_{jq})^2 \geq \sum_{p,q=1}^2 \frac{|A_p^{ij}||B_q^{ij}|}{|X_i||X_j|} d(A_p^{ij}, B_q^{ij})^2 \geq d(X_i, X_j)^2 + \varepsilon^4 \text{ với mọi } (i, j) \in I. (*)$$

Nhân hai vế của đánh giá (*) cho $\frac{|X_i||X_j|}{n^2}$ ta thu được :

$$\sum_{p=1}^{r_i} \sum_{q=1}^{r_j} \frac{|X_{ip}||X_{jq}|}{n^2} d(X_{ip}, X_{jq})^2 \geq \frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2 + \varepsilon^4 \frac{|X_i||X_j|}{n^2} \text{ với mọi } (i, j) \in I$$

Lấy tổng tất cả các bộ $(i, j) \in \{1, 2, \dots, m\} \times \{1, 2, \dots, m\}$ và chia thành hai loại $(i, j) \in I$ và $(i, j) \in I^C$ và sử dụng kết quả 2) ta thu được đánh giá sau:

$$\sum_{i,j=1}^m \sum_{p=1}^{r_i} \sum_{q=1}^{r_j} \frac{|X_{ip}| |X_{jq}|}{n^2} d(X_{ip}, X_{jq})^2 \geq \underbrace{\sum_{i,j=1}^m \frac{|X_i| |X_j|}{n^2} d(X_i, X_j)^2}_{\alpha^2} + \varepsilon^4 \sum_{(i,j) \in I} \frac{|X_i| |X_j|}{n^2} \geq \alpha^2 + \varepsilon^5$$

(Lưu ý: $X_1, X_2, \dots, X_m$ là một phân hoạch không ε -đều của $V(G)$ nên $\sum_{(i,j) \in I} \frac{|X_i| |X_j|}{n^2} \geq \varepsilon$)

Từ đây ta suy ra điều phải chứng minh.

Kết quả 5: (Szemerédi)

Cho trước số $\varepsilon > 0$. Luôn tồn tại số nguyên dương $k = k(\varepsilon)$ sao cho với mọi đồ thị G ta luôn tìm được một phân hoạch ε -đều của $V(G)$ thành tối đa k tập.

Chứng minh:

Đầu tiên ta xét phân hoạch tầm thường của $V(G)$ thành chính nó. Nếu phân hoạch này là ε -đều ta chỉ chọn $k(\varepsilon) = 1$ và kết thúc bài toán. Trong trường hợp ngược lại theo kết quả 4) ta có thể làm mịn phân hoạch này thành 4 tập sao cho trung bình bình phương mật độ tính theo phân hoạch này tăng tối thiểu là ε^5 so với khi tính theo phân hoạch ban đầu... Tiếp tục quá trình tính toán như trên, nếu tại bước thứ k ta phân hoạch $V(G)$ thành m tập thì tại bước thứ $k+1$ ta phân hoạch $V(G)$ thành tối đa $m2^{2^m} \leq 2^{2^m}$ tập và trung bình bình phương mật độ tăng lên tối đa là $k\varepsilon^5$. Quá trình trình phải kết thúc sau hữu hạn bước, có nghĩa là sau một số hữu hạn bước thì phân hoạch mà ta thu được phải là ε -đều, nếu ngược lại thì ta xét tại bước thứ $k \geq \frac{2}{\varepsilon^5}$ thì áp dụng kết quả 4) và một lập luận qui nạp đơn giản ta dễ dàng nhận thấy giá trị của trung bình bình phương mật độ tăng tối thiểu 2 đơn vị so với giá trị khi tính theo phân hoạch tầm thường chỉ gồm $V(G)$, điều này mâu thuẫn với việc giá trị của trung bình bình phương nằm giữa 0 và 1.

Tài liệu tham khảo:

Timothy Gowers, Topic in Combinatorics