Mục lục

Lời mở đầu

Hà Nội, tháng 6 năm 2020

Sinh viên thực hiện

Lai Đức Thắng

Chương 1

Vĩnh thức

1.1 Khái niệm

Giả sử $M = (m_{ij})$ là ma trận số thức $n \times n$. Khi đó vĩnh thức (Permanents) của ma trận M (kí hiệu là per):

$$perM = \sum_{\sigma \in S_n} m_{1\sigma(1)} m_{2\sigma(2)} ... m_{n\sigma(n)}$$

trong đó S_n là tập tất cả các hoán vị của $\{1, 2, ..., n\}$.

1.2 Phỏng đoán của Van der Waerden

Khác với định thức, có thể tính toán nhanh chóng (sử dụng phép khử Gaussian), việc tính toán với vĩnh thức là khá khó khăn. Một vài nghiên cứu gần đây về vĩnh thức xem xét về xấp xỉ và giới hạn của giá trị này. Trong nội dung bài báo cáo này ta xem xét đến một định lý nổi tiếng về vĩnh thức và chứng minh của nó. Một ma trận $M = (m_{ij})$ được gọi là *ngẫu nhiên kép* (*doubly stochastic*) nếu các phần tử của ma trận là các số thực không âm sao cho tổng theo mỗi hàng hoặc mỗi cột bằng 1. Năm 1926 Bartel L. Vander Waerden đưa ra phỏng đoán:

$$\text{per } M \geq \frac{n!}{n^n}$$

đúng với mọi ma trận *ngẫu nhiên kép* $n \times n$. Dấu "="xảy ra khi và chỉ khi $M = (m_{ij})$, trong đó $m_{ij} = \frac{1}{n}$ với mọi i và j Phỏng đoán này chưa được giải quyết trong hơn 50 năm, cho đến khi được xác nhận bởi G. P. Egorchev và D. I. Falikman vào băn 1981. Sau đó, vào năm 2007, Leonid Gurvits đưa ra một chứng minh ngắn gọn và hoàn toàn khác biệt. Trước hết ta sẽ phát biểu lại định lý.

Định lý 1. *Đặt $M = (m_{ij})$ là ma trận ngẫu nhiên kép $n \times n$. Khi đó*

$$per\ M \geq \frac{n!}{n^n} \tag{1.1}$$

Dấu "="xảy ra khi và chỉ khi $m_{ij} = \frac{1}{n}$ với mọi i và j.

Đầu tiên ta sẽ chuyển ma trận về dạng đa thức. Với mọi ma trận $n \times n$ $M = (m_{ij})$, ta xây dựng một đa thức $p_M(x) \in R_{[x_1,...,x_n]}$,

$$p_M(x) = p_M(x_1, .., x_n) := \prod_{i=1}^{n} \Big(\sum_{j=1}^{n} m_{ij} x_j \Big).$$

Tiếp theo ta định nghĩa đạo hàm của $p_M(x) \in R_{[x_1,...,x_n]}$ theo biến x_n:

$$p'(x_1, .., x_{n-1}) := \frac{\partial p(x)}{\partial x_n} \mid_{x_n=0}$$

Quan sát rằng p là đa thức thuần nhất bậc n với n biến, khi đó p' là đa thức thuần nhất bậc $n-1$ với $n-1$ biến. Tổng quát, với $i = 0, 1, .., n$

$$q_i(x_1, .., x_i) := \frac{\partial^{n-i} p(x)}{\partial x_n ... \partial x_{i+1}} \mid_{x_n = x_{n-1} = ... = x_{i+1} = 0}$$

Từ công thức trên ta nhận được một dãy $(q_n, q_{n-1}, ..., q_0)$, trong đó $q_n = p$ và $q_{i-1} = q_i'$ với $1 \leq i \leq n$ và q_0 là hệ số của $x_1 x_2 ... x_n$ trong đa thức p. Thêm nữa, nếu p là đa thức thuần nhất bậc n, thì q_i là đa thức thuần nhất bạc q_i. Xét dãy sinh ra bởi đa thức $p_M(x)$,

$$p_M(x) = q_n, ..., q_i, ..., q_0$$

Ta suy ra hai điều quan trọng sau đây:

A. *per M* là hệ số của $x_1 x_2 ... x_n$ trong q_n, do đó $q_0 = \text{per } M$.

B. Với $i = 1, .., n$ ta có

$$deg_i q_i \leq min\{i, \lambda_M(i)\}, \tag{1.2}$$

trong đó $deg_i q_i$ kí hiệu là bậc của x_j trong $q_i(x_1, ..., x_n)$ và $\lambda_M(i)$ là số các giá trị khác 0 trong cột thứ i của ma trận M.

Thật vậy, ta có $deg_i q_i \leq i$ vì q_i là đa thức thuần nhất bậc i, trong khi $deg_i q_i \leq deg_i q_n \leq \lambda_M(i)$ là hiển nhiên theo định nghĩa của $p_M(x)$.

Sau đây là ý tưởng chính của chứng minh: Ta liên kết một tham số với mỗi đa thức p và xác định một cận dưới cho tham số đó khi truyền từ p sang p'.

Ta kí hiệu $\mathbb{R}_+$ là tập các số thực không âm và $p(x) \in \mathbb{R}_{+[x_1,...,x_n]}$ là đa thức trong đó các hệ số của $p(x)$ là không âm. Với số thức $z \in \mathbb{C}$, đặt $Re(z)$ và $Im(z)$ lần lượt là phần thực và phần ảo của z. Đặt $\mathbb{C}_+ = \{z \in \mathbb{C} : Re(z) \geq 0\}$ và $\mathbb{C}_{++} = \{z \in \mathbb{C} : Re(z) > 0\}$. Kí hiệu này mở rộng với $\mathbb{R}_+^n$ và $\mathbb{C}_+^n$. Ví dụ, $z = (z_1, ..., z_n) \in \mathbb{C}_{++}^n$ đúng nếu $Re(z_i) > 0$ với mọi i.

Với mọi đa thức $p(x) \in \mathbb{R}_{+[x_1,...,x_n]}$ ta định nghĩa **capacity** của p, kí hiệu $cap(p)$ bởi:

$$\text{cap}(p) := \inf \{p(x) : x \in \mathbb{R}_+^n, \prod_{i=1}^{n} x_i = 1\}$$

Đặc biệt $\text{cap}(p) \geq 0$ vì p chỉ có các hệ số không âm, và nếu p là hằng số ($p(x) \equiv c$) thì $\text{cap}(p) = c$ Ngoài ra ta cần hàm $g : \mathbb{N}_0 \to \mathbb{R}$ với $g(0) := 1$ và

$$g(k) := \Big(\frac{k-1}{k}\Big)^{k-1} \text{ với } k \geq 1.$$

Sử dụng bất đẳng thức $1 + x \leq e^x$ 2 lần, ta được

$$\frac{g(k+1)}{g(k)} = \frac{k}{k+1}\Big(\frac{k^2}{k^2-1}\Big)^{k-1} < e^{-\frac{1}{k+1}} e^{\frac{1}{k^2-1}} = 1$$

với $k \geq 1$. Do đó, g là hàm không tăng, $g(0) = g(1) > g(2) > ...$.

Ta gọi đa thức $p(x) \in \mathbb{R}_{[x_1,...,x_n]}$ là *H-ổn định* nếu đa thức này không có nghiệm trên C_{++}^n.

Mệnh đề Gurvits *Nếu $p(x) \in \mathbb{R}_{+[x_1,...,x_n]}$ là H-ổn định và thuần nhất bậc n, khi đó hoặc $p' \equiv 0$ hoặc p' là H-ổn định và thuần nhất bậc $n-1$. Trong cả hai trường hợp*

$$\text{cap}(p') \geq \text{cap}(p).g(\deg_n p) \tag{1.3}$$

Chứng minh Định lý 1. Đặt $M = (m_{ij})$ là ma trận *ngẫu nhiên kép* $n \times n$. Ta đã biết $p_M(x)$ là đa thức thuần nhất bậc n. Khi đó hai khẳng định dưới đây là đúng:

Khẳng định 1. *$p_M(x)$ là H-ổn định*

Bằng phản chứng, giả sử x là nghiệm của $p_M(x)$. Từ $p_M(x) = \prod_{i=1}^{n}(\sum_{j=1}^{n} m_{ij}x_j) = 0$ suy ra $\sum_{j=1}^{n} m_{ij}x_j = 0$ nên $\sum_{j=1}^{n} m_{ij}Re(x_j) = 0$. Điều này trái với giả thiết $x \in \mathbb{C}^n_{++}$, vì $m_{il} > 0$ với một vài giá trị l.

Khẳng định 2. *$cap(p_M) = 1$*

Chứng minh. Trước tiên ta nhắc lại bất đẳng thức $AM-GM$: Với $a_1, ..., a_n, p_1, .., p_n \in \mathbb{R}_+$ thoả mãn $\sum_{i=1}^{n} p_i = 1$ ta có

$$\sum_{i=1}^{n} p_i a_i \geq a_1^{p_1} ... a_n^{p_n}.$$

Chọn bất kì giá trị $x \in \mathbb{R}^n_+$ với $\prod_{j=1}^{n} x_j = 1$. Áp dụng bất đẳng thức AM-GM:

$$\begin{aligned} p_M(x) = \prod_{i=1}^{n}\Big(\sum_{j=1}^{n} m_{ij}x_j\Big) &\geq \prod_{i=1}^{n}\prod_{j=1}^{n} x_j^{m_{ij}} \\ &= \prod_{j=1}^{n}\prod_{i=1}^{n} x_j^{m_{ij}} = \prod_{j=1}^{n} x_j^{\sum_{i=1}^{n} m_{ij}} \\ &= \prod_{j=1}^{n} x_j = 1 \end{aligned}$$

do đó $cap(p_M) \geq 1$.

Mặt khác,

$$p_M(1, 1, ..., 1) = \prod_{i=1}^{n}\Big(\sum_{j=1}^{n} m_{ij}\Big) = \prod_{i=1}^{n} 1 = 1, \tag{1.4}$$

□

Vì $p_M(x)$ là *H-ổn định*, ta có thể áp dụng mệnh đề (1.3) nhiều lần để đưa ra kết luận mọi đa thức q_i đều là *H-ổn định*, sao cho với mỗi giá trị của i:

$$cap(q_{i-1}) \geq cap(q_i)g(deg_i q_i) \geq cap(q_i)g(min\{i, \lambda_M(i)\}), \tag{1.5}$$

trong đó bất đẳng thức thứ hai được suy ra từ (1.2) với g là hàm giảm.

Lặp lại (1.5) bắt đầu với $cap(p_M) = 1$, ta có:

$$\begin{aligned} perM = q_0 &\geq \prod_{i=1}^{n} g(min\{i, \lambda_M(i)\}) \\ &\geq \prod_{i=1}^{n} g(i) = \prod_{i=1}^{n} \left(\frac{i-1}{i}\right)^{i-1} = \prod_{i=1}^{n} i\frac{(i-1)^{i-1}}{i^i} = \frac{n!}{n^n} \end{aligned}$$

□

Chương 2

Lý thuyết thông tin

2.1 Khái niệm

2.2 Entropy

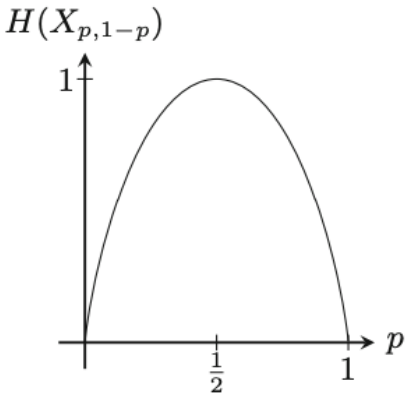
Giả sử X là một biến ngẫu nhiên nhận các giá trị $\{a_1, ..., a_n\}$ với xác suất $Pr(X = a_i) = p_i$. Nó giúp ta suy nghĩ về việc coi X là một phép thử với các khả năng là a_i, giống như việc tung một xúc xắc với các khả năng là số chấm trên xúc xắc đó $1, 2, 3, 4, 5, 6$. Vậy lượng thông tin (về trung bình) nhận được từ phép thử trên là bao nhiêu?
Với biến ngẫu nhiên rời rạc X nhận các giá trị $\mathscr{X} = \{x_1, ..., x_n\}$ và hàm phân phối xác suất $Pr(X)$ thì Entropy của X là:

$$H(X) = \sum_{i=1}^{n} Pr(x_i) \log_b \frac{1}{Pr(x_i)},$$

trong đó ta quy ước $0.log(\frac{1}{0}) = 0$ (giả định rằng tổng chỉ được lấy trên các phần tử $x \in \mathscr{X}$ sao cho $Pr(X = x) > 0$). Khi đó *giá*(*support*) của biến ngẫu nhiên X là: *supp* $X := \{a : Pr(X = a) > 0\}$. Biểu thức trên cũng có thể được viết lại thành:

$$H(X) = -\sum_{i=1}^{n} Pr(x_i) \log_2 Pr(x_i)$$

Đôi lúc để ký hiệu tiện lợi và dễ nhìn hơn chúng ta có thể viết lại công thức Entropy



Hình 2.1: x

với vector xác suất $p = (p_i, ..., p_n)$ với $p_i = Pr(X = x_i)$. Khi đó:

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

Xét một ví dụ, nếu X là sự kiện tung một đồng xu với $Pr(X = \text{ngửa}) = p$, khi đó theo công thức của Shannon cho ta hàm $H(X_{p,1-p}) = -p \log_2 p - (1-p) \log_2 (1-p)$ (hình 2.1).

2.3 Entropy có điều kiện

Giả sử X và Y là hai biến ngẫu nhiên nhận các giá trị lần lượt $\{a_1, .., a_m\}$ và $\{b_1, .., b_n\}$. Entropy hợp (*joint entropy*) của 2 biến ngẫu nhiên X và Y là:

$$H(X, Y) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

Nếu $p(b_j|a_i) := Pr(Y = b_j|X = a_i)$ là xác suất có điều kiện của b_j khi biết a_i. Khi đó entropy có điều kiện của Ynếu ta biết kết quả của X là a_i là:

$$H(Y|a_i) := -\sum_{i=1}^{n} p(b_j|a_i) \log_2 p(b_j|a_i)$$

Nếu lấy giá trị kì vọng của biểu thức trên với tất cả các khả năng của X, ta thu được:

$$H(Y|X) := \sum_{i=1}^{n} p(a_i) H(Y|a_i)$$

chính là entropy có điều kiện của Y khi biết X Sau đây ta sẽ xem một số mối quan hệ giữa các đại lượng ở trên.

Mệnh đề 1. *Giả sử X và Y là hai biến ngẫu nhiên nhận các giá trị trong tập $\mathscr{X}$ và $\mathscr{Y}$. Ta có một số kết quả sau:*

(A) $H(X) \leq \log_2(|suppX|)$.

(B) $H(X,Y) = H(X) + H(Y|X)$, và tổng quát ta có $H(X_1,...,X_n) = H(X_1) + H(X_2|X1) + ... + H(X_n|X_1,...,X_{n-1})$.

(C) $H(X,Y) \leq H(X) + H(Y|X)$. Dấu "="xảy ra khi và chỉ khi X và Y là độc lập.

(D) $H(X,Y) \leq H(X)$.

(E) Nếu *supp* X được chia thành d tập $E_1,..,E_d$ sao cho $E_j := \{a \in suppX : |supp(Y|a)| = j\}$ thì:

$$H(Y|X) \leq \sum_{j=1}^{d} Pr(X \in E_j) \log_2 j.$$

Trước khi đi vào chứng minh, ta nhắc lại bất đẳng thức $AM - GM$:

$$a_1^{p_1}.....a_n^{p_n} \leq p_1a_1.....p_na_n$$

Chứng minh mệnh đề 1:

(A) Không mất tính tổng quát, giả sử $p_i > 0\ \forall i$. Áp dụng bất đẳng thức **AM-GM**: $a_1^{p_1}.....a_n^{p_n} \leq p_1a_1.....p_na_n$: Đặt $a_i = \frac{1}{p_i}$ và lấy log 2 vế, ta được:

$$H(X) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} \leq \log_2 \Big(\sum_{i=1}^{n} p_i \frac{1}{p_i} \Big) = \log_2 n.$$

Dấu "="xảy ra khi và chỉ khi $p_1 = ... = p_n = \frac{1}{n}$.

(B)

$$\begin{aligned}
H(X,Y) &= -\sum_{i,j} p(a_i,b_j)\log_2 p(a_i,b_j) \\
&= -\sum_{i,j} p(a_i,b_j)\log_2 p(a_i)p(b_j|a_i) \\
&= -\sum_{i,j} p(a_i,b_j)[\log_2 p(a_i) + \log_2 p(b_j|a_i)] \\
&= -\sum_{i,j} p(a_i,b_j)\log_2 p(a_i) - \sum_{i,j} p(a_i,b_j)\log_2 p(b_j|a_i) \\
&= -\sum_{i,j} p(a_i,b_j)\log_2 p(a_i) - \sum_{i,j} p(a_i).p(b_j|a_i)\log_2 p(b_j|a_i) \\
&= -\sum_{i=1}^{m} p(a_i)\log_2 p(a_i)\sum_{j=1}^{n} p(b_j|a_i) + H(Y|X) \\
&= -\sum_{i=1}^{m} p(a_i)\log_2 p(a_i) + H(Y|X) = H(X) + H(Y|X).
\end{aligned}$$

(Do $\sum_{j=1}^{n} p(b_j|a_i) = \sum_{j=1}^{n} \frac{P(b_j.a_i)}{P(a_i)} = \frac{1}{P(a_i)}\sum_{j=1}^{n} P(b_j.a_i) = \frac{1}{P(a_i)}P(a_i) = 1$)

(C) Vì $Pr(X=x) = \sum_{y\in\mathscr{Y}} Pr(X=x,Y=y)$,

$H(X)+H(Y) = -\sum_{x,y} Pr(X=x,Y=y)\log(Pr(X=x).Pr(Y=y))$.sử dụng bất đẳng thức ...., ta có:

$$\begin{aligned}
H(X,Y) - (H(X)+H(Y)) &= \sum_{x,y} Pr(X=x,Y=y)\log\left(\frac{Pr(X=x).Pr(Y=y)}{Pr(X=x,Y=y)}\right) \\
&\leq \log\left(\sum_{x,y} Pr(X=x).Pr(Y=y)\right) \\
&= \log 1 = 0
\end{aligned}$$

Dâu "="xảy ra khi và chỉ khi X,Y độc lập với nhau.

(D) Theo (B) và (C):

$$H(Y|X) - H(X) = H(X,Y) - H(Y) - H(X) \leq 0,$$

Dấu "="xảy ra khi và chỉ khi X,Y độc lập với nhau.

(E) Ta có *supp* $(Y|a) = \{b : Pr(Y=b|X=a) > 0\}$ .Vì $Pr(Y=b|X=a)$ là 1 biến ngẫu nhiên trên tập *supp* $(Y|a)$ (Do $\sum_{i=1}^{m} Pr(Y=b_i|a) = 1$) Tiến hành chia tập *supp* X thành các tập con E_j theo giả thuyết và sử dụng kết quả từ **(A)**, ta có:

$$\begin{aligned} H(X,Y) &= \sum_{i=1}^{m} p(a_i) H(Y|a_i) \\ &= \sum_{j=1}^{d} \sum_{a \in E_j} p(a) H(Y|a) \\ &\leq \sum_{j=1}^{d} \sum_{a \in E_j} p(a) \log_2 j \\ &= \sum_{j=1}^{d} Pr(X \in E_j) \log_2 j. \end{aligned}$$

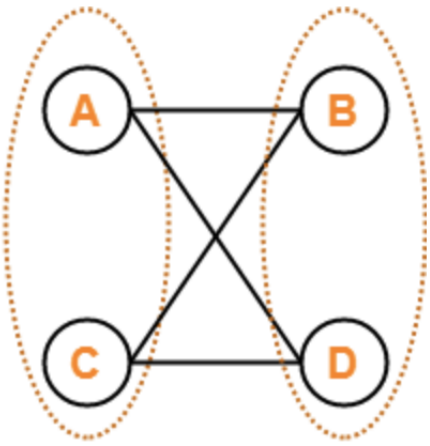
Chương 3

Entropy và bài toán Vĩnh thức

3.1 Đồ thị hai phía

3.1.1 Khái niệm

Một đồ thị đơn vô hướng $G := (V, E)$ được gọi là đồ thị hai phía nếu tồn tại một phân hoạch của tập đỉnh $V = V_1 \cup V_2$ sao cho V_1 và V_2 là các tập độc lập và thoả mãn: bất kỳ cạnh nào thuộc tập cạnh E của đồ thị cũng nối một đỉnh $\in V_1$ với một đỉnh $\in V_2$. Ta thường viết $G := (V_1 \cup V_2, E)$ để ký hiệu một đồ thị hai phía với các phân hoạch V_1 và V_2. Nếu $|V_1| = |V_2|$ thì G được gọi là đồ thị hai phía cân bằng.
Ví dụ về đồ thị hai phía:



Đồ thị trên là một đồ thị hai phía vì:

- Các đỉnh của đồ thị có thể phân tách thành hai tập $V_1 = \{A, C\}$ và $V_2 = \{B, D\}$
- Các đỉnh của tập V_1 chỉ nối với các đỉnh của tập V_2 và ngược lại
- Các đỉnh trong cùng một tập không kề với nhau
- $|V_1| = |V_2| = 2$ nên đây là đồ thị hai phía cân bằng

Một **cặp ghép** (matching) trong một đồ thị $G := (V, E)$ là tập các cạnh $M \subseteq E$ đôi một không có điểm chung. Một đỉnh của đồ thị được gọi là đã ghép (matched) nếu nó là đầu mút của một cạnh trong M. Ngược lại, ta gọi đỉnh đó là đỉnh chưa ghép (unmatched). Một cặp ghép được gọi là **hoàn hảo** (perfect matching) nếu mọi đỉnh của đồ thị đều đã được ghép.

3.1.2 Tính chất

Một đồ thị hai phía có thể được mô tả qua những cách sau:

- Một đồ thị là hai phía khi và chỉ khi nó không chứa chu trình lẻ.
- Một đồ thị là hai phía khi và chỉ khi có thể tô màu nó bằng hai màu.
- Phổ của đồ thị là đối xứng khi và chỉ khi đó là đồ thị hai phía

3.1.3 Ứng dụng

Đồ thị hai phía thường được dùng để mô hình các bài toán ghép cặp (matching problem). Một ví dụ bài toán phân công công việc. Giả sử ta có một nhóm người P và một tập công việc J, trong đó không phải ai cũng hợp với mọi công việc. Ta có thể mô hình bài toán bằng một đồ thị với tập đỉnh là $P + J$. Nếu người p_i có thể làm công việc j_i, đồ thị sẽ có một cạnh nối giữa p_i và j_i. Định lý hôn nhân cung cấp một đặc điểm của đồ thị hai phía: tồn tại cặp ghép hoàn hảo (perfect matching).
Đồ thị hai phía được sử dụng trong lý thuyết mã hóa (coding theory) hiện đại, đặc biệt khi giải mã các codeword nhận được từ kênh. Đồ thị nhân tử (factor graph) và đồ thị Tanner là các ví dụ.

3.2 Định lý Brégman

Mô hình bài toán

Xét ma trận $n \times n$ $M = (m_{ij})$ với các phần tử $\{0, 1\}$. Ta liên kết M với một đồ thị hai phía $G_M = (U \cup V, E)$, trong đó $U = \{u_1, .., u_n\}$, $V = \{v_1, .., v_n\}$ thoả mãn:

$$u_i v_j \in E \Longleftrightarrow m_{ij} = 1$$

Ngược lại, mọi đồ thị hai phía G với $n+n$ nút cho ta một ma trận $\{0, 1\}$ có kích thước $n \times n$ với $G = G_M$. Xét biểu thức vĩnh thức của ma trận M:

$$perM = \sum_{\sigma \in S_n} m_{1\sigma(1)} m_{2\sigma(2)} ... m_{n\sigma(n)}$$

Khi đó, mỗi thành phần $m_{1\sigma(1)} m_{2\sigma(2)} ... m_{n\sigma(n)}$ nhận giá trị 0 hoặc 1; bằng 1 khi và chỉ khi tập các cạnh $\{u_1 v_{\sigma(1)}, ..., u_n v_{\sigma(n)}\}$ là một *ghép cặp hoàn hảo (perfect matching)* của G_M và ngược lại. Do đó số lượng *ghép cặp hoàn hảo* $m(G_M)$ chính là giá trị của vĩnh thức, hay $perM = m(G_M)$ Mối liên hệ giữa $G \leftrightarrow M_G$ đã minh hoạ cho một nghiên cứu về vĩnh thức. Một trong những vấn đề khó khăn đầu tiên là một phỏng được đưa ra bởi Henryk Minc vào năm 1967: Giả sử ma trận $M - \{0/1\}$ có tổng theo hàng $d_1, ..., d_n$ (tương đương với các đỉnh $u_1, ..u_n$ có bậc $d_1, ..d_n$), khi đó:

$$perM \leq \prod_{i=1}^{n} (d_i!)^{1/d_i}.$$

Phỏng đoán trên của Minc được Lev M. Brégman chứng minh vào năm 1973. Một vài năm sau Alexander Schrijver đưa ra một chứng minh khác ngắn hơn. Tuy nhiên, trong nội dung báo cáo này, em xin trình bày lại chứng minh của Jaikumar Radhakrishnan, sử dụng *Entropy từ lý thuyết thông tin.* Trước tiên ta nhắc lại định lý Brégman :

Định lý 1. Đặt $M = (m_{ij})$ là ma trận $n \times n$ chỉ chứa hai giá trị $\{0, 1\}$ và đặt $d_1, ..., d_n$ là tổng các hàng của ma trận M, hay $d_i = \sum_{j=1}^{n} m_{ij}$. Khi đó:

$$perM \leq \prod_{i=1}^{n} (d_i!)^{1/d_i}.$$

Chứng minh. Xét $G = (U \cup V, E)$ là đồ thị hai phía tương ứng với ma trận M, trong đó d_i là bậc tương ứng của các đỉnh u_i, và kí hiệu $\mathfrak{S}$ là tập các *perfect matching* của G. Vì $perM = m(G) = |\mathfrak{S}|$ nên thay vì tìm cận trên cho $perM$ như định lý 1, ta sẽ tìm cận trên cho $|\mathfrak{S}|$. Giả sử $|\mathfrak{S}| \neq 0$ và mỗi $\sigma \in \mathfrak{S}$ là một hoán vị tương ứng $\sigma(1)\sigma(2)....\sigma(n)$ của các chỉ số. Vì vậy, tương ứng với mỗi giá trị $u_i \in U$ là một giá trị $v_{\sigma(i)} \in U$ theo phép song ánh σ. Ý tưởng ban đầu là chọn σ một cách ngẫu nhiên và xét biến ngẫu nhiên $X = (X_1, X_2, .., X_n) = (\sigma(1), \sigma(2), ..., \sigma(n))$.

Theo mệnh đề **(A)**,

$$H(\sigma(1), \sigma(2), ..., \sigma(n)) = \log_2(|\mathfrak{S}|)$$

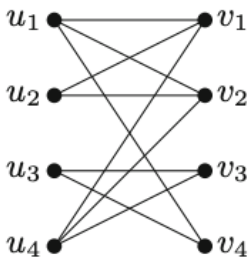
Do đó chỉ cần chỉ ra

$$H(\sigma(1), ..., \sigma(n)) \leq \log_2 \Big(\prod_{i=1}^{n} (d_i!)^{1/d_i} \Big) = \sum_{i=1}^{n} \frac{1}{d_i} \log_2(d_i!). \tag{3.1}$$

Tiếp theo, sử dụng mệnh đề **(B)**, ta có

$$H(\sigma(1), \sigma(2), ..., \sigma(n)) = \sum_{i=1}^{n} H(\sigma(i)|\sigma(1), \sigma(2), ..., \sigma(i-1)) \tag{3.2}$$

Vậy biểu thức entropy có điều kiện ở trên có ý nghĩa gì ? Nó đo mức độ không chắc chắc (*uncertainty*) về đỉnh được ghép với u_i sau khi các tập các đỉnh ghép với $u_1, ..., u_{i-1}$ được xác định. Cụ thể, *giá* (*support*) của biến ngẫu nhiên $\sigma(i)$ khi biết $(\sigma(1), ..., \sigma(i-1))$ nằm trong tập chỉ số của lân cận u_i mà chưa được ghép với một trong các đỉnh $u_1, ..., u_{i-1}$. Lấy ví dụ, xét đồ thị hai phía (Hình 3.1), có $|\mathfrak{S}| = 4$. Vì mọi hoán vị trong $\mathfrak{S}$ là như nhau nên $H(\sigma(1), ..., \sigma(4)) = \log_2 4 = 2$. Khi đó, $H(\sigma(1)) = -\frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{2}\log_2\frac{1}{2} = \frac{3}{2}$. Tiếp theo, ta sẽ xác định entropy có điều kiện $H(\sigma(2)|\sigma(1))$: Với $\sigma(1) = 1$ ta có $H(\sigma(2)|1) = 0$ vì $\sigma(2) = 2$ là duy nhất; tương tự cho $H(\sigma(2)|2) = 0$ (vì $\sigma(2) = 1$ là duy nhất), nhưng với $\sigma(1) = 4$ thì $H(\sigma(2)|4) = 1$ vì có hai khả năng cho $\sigma(2)$ là $\sigma(2) = 1$ và $\sigma(2) = 2$. Lấy giá trị kì vọng, ta được $H(\sigma(2)|\sigma(1)) = \frac{1}{2}.1 = \frac{1}{2}$. Biểu thức entropy có điều kiện tiếp theo $H(\sigma(3)|\sigma(1), \sigma(2))$ và $H(\sigma(4)|\sigma(1), \sigma(2), \sigma(3))$ đều bằng 0 vì các giá trị của $\sigma(1), \sigma(2), \sigma(3)$ được xác định đều là duy nhất. Lấy tổng các entropy có điều kiện này lại, ta có $H(\sigma(1)) + H(\sigma(2)|\sigma(1)) + H(\sigma(3)|\sigma(1), \sigma(2)) + H(\sigma(4)|\sigma(1), \sigma(2), \sigma(3)) = \frac{3}{2} + \frac{1}{2} + 0 + 0 = 2$, thoả mãn mệnh đề **(B)**. Ý tưởng của Radhakrishnan là xét các đỉnh $u_1, u_2, ..., u_n$ theo một *thứ tự ngẫu nhiên* $\tau \in S_n$, với xác suất là như nhau và bằng



$$\mathfrak{S} = \{1243, 2143, 4132, 4231\}$$

Hình 3.1: Đồ thị hai phía với $\mathfrak{S}$ là tập các ghép cặp hoàn hảo

$\frac{1}{n!}$, và lấy giá trị kì vọng của các entropy. Nói cách khác, ta xét các cặp *matching* theo thứ tự $\tau_1, \tau_2, .., \tau_n$. Xét τ cố định, khi đó $k_i = \tau_i^{-1}$ được hiểu là vị trí của u_i theo thứ tự ngẫu nhiên τ là k_i. Khi đó, biểu thức (3.2) trở thành:

$$H(\sigma(1), ..., \sigma(n)) = \sum_{i=1}^{n} H\Big(\sigma(i)|\sigma(\tau_1), ..., \sigma(\tau_{k_i-1})\Big)$$

Khi đó

$$H(\sigma(1), ..., \sigma(n)) = \frac{1}{n!}\sum_{i=1}^{n} H\Big(\sigma(i)|\sigma(\tau_1), ..., \sigma(\tau_{k_i-1})\Big)$$

Xét biểu thức

$$H\Big(\sigma(i)|\sigma(\tau_1), ..., \sigma(\tau_{k_i-1})\Big) \tag{3.3}$$

Để tìm cận trên cho (3.3), ta sẽ sử dụng mệnh đề **(C)**, áp dụng với biến ngẫu nhiên $X = \Big(\sigma(\tau_1), ..., \sigma(\tau_{k_i-1})\Big)$ và $Y = \sigma(i)$. Với τ cố định $\in S_n$ và $\sigma \in \mathfrak{S}$ đặt $N_i(\sigma, \tau)$ là số các giá trị $k \in [n]$ sao cho $u_i v_k \in E(G)$ và k $\notin \Big(\sigma(\tau_1), ..., \sigma(\tau_{k_i-1})\Big)$ (nói cách khác, $N_i(\sigma, \tau)$ là số khả năng còn lại cho $\sigma(i)$ khi đã biết $(\sigma(\tau_1), ..., \sigma(\tau_{k_i-1}))$. Vì $deg(u_i) = d_i$ đồng thời u_i phải được ghép cặp trong σ nên $1 \leq N_i(\sigma, \tau) \leq d_i$ với mọi $\sigma \in \mathfrak{S}$. Tiến hành chia tập $suppX$ thành các tập con $E_{i,j}^{(\tau)}$ sao cho :

$$\Big(\sigma(\tau_1), ..., \sigma(\tau_{k_i-1})\Big) \in E_{i,j}^{(\tau)} <=> N_i(\sigma, \tau) = j, \text{ với } 1 \leq j \leq d_i$$

Coi $N_i(\sigma, \tau)$ là một biến ngẫu nhiên trên $\mathfrak{S}$, ta có:

$$Pr(X \in E_{i,j}^{(\tau)}) = Pr(N_i(\sigma, \tau) = j)$$

Từ mệnh đề **(C)**, với τ cố định:

$$\begin{aligned} H(\sigma(i)|\sigma(\tau_1),...,\sigma(\tau_{k_i-1})) &\leq \sum_{j=1}^{d_i} Pr(|N_i(\sigma,\tau)|=j)\log_2 j \\ &= \sum_{j=1}^{d_i}\log_2 j \sum_{\sigma\in\mathfrak{S}} P(\sigma).P(N_i(\sigma,\tau)=j|\sigma) \text{ (Do } \sigma\in\mathfrak{S} \text{ là 1 nhóm đầy đủ)} \\ &= \sum_{j=1}^{d_i}\log_2 j \sum_{\sigma\in\mathfrak{S}} \frac{P(N_i(\sigma,\tau)=j|\sigma)}{|\mathfrak{S}|} \end{aligned}$$

Kết hợp với ...

$$H(\sigma(1),...,\sigma(n)) \leq \sum_{j=1}^{d_i}\log_2 j\left(\frac{1}{n!|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}}\sum_{\tau\in S_n} P(N_i(\sigma,\tau)=j|\sigma)\right) \tag{3.4}$$

Xét $\sum_{\tau\in S_n} P(N_i(\sigma,\tau)=j|\sigma)$:
Với mỗi σ cố định $\in\mathfrak{S}$, $N_i(\sigma,\tau)$ nhận các giá trị từ 1 tới d_i với xác suất như nhau là $\frac{1}{d_i}$, vì $N_i(\sigma,\tau)$ chỉ phụ thuộc vào vị trí của $\sigma(i)$ trong hoán vị τ (do σ đã cố định), tương ứng với các đỉnh k thỏa mãn $u_iv_k\in E(G)$ (Nếu $\sigma(i)$ là lân cận gần nhất của i theo thứ tự của hoán vị (giả sử là τ_1). Khi đó $N_i(\sigma,\tau_1)=d_i$ và $P(N_i(\sigma,\tau_1)=d_i|\sigma)=\frac{1}{d_i}$; Nếu $\sigma(i)$ là lân cận gần thứ hai của i theo thứ tự của hoán vị (giả sử τ_2). Khi đó $N_i(\sigma,\tau_2)=d_i-1$ và $P(N_i(\sigma,\tau_2)=d_i-1|\sigma)=\frac{1}{d_i}$,...tương tự với n! hoán vị của S_n .Do đó, giá trị của biểu thức trên bằng:

$$\sum_{\tau\in S_n} P(N_i(\sigma,\tau)=j|\sigma) = n!.\frac{1}{d_i}=\frac{n!}{d_i}$$

Và biểu thức (3.2) trở thành:

$$\begin{aligned} H(\sigma(1),\sigma(2),....,\sigma(n)) &= \sum_{i=1}^{n} H(\sigma(i)|\sigma(1),\sigma(2),....,\sigma(i-1)) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{d_i}\log_2 j\left(\frac{1}{n!|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}}\frac{n!}{d_i}\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{d_i}\log_2 j\left(\frac{1}{n!|\mathfrak{S}|}.|\mathfrak{S}|.\frac{n!}{d_i}\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{d_i}\log_2 j.\frac{1}{d_i} = \sum_{i=1}^{n}\frac{\log_2 d_i!}{d_i}. \quad \square \end{aligned}$$

Chương 4

Hình vuông Latin

4.1 Khái niệm

Hình vuông latin là một trong số những hình tổ hợp lâu đời nhất, có những nghiên cứu rõ ràng từ thời cổ đại. Để có được hình vuông Latin, ta phải điền vào n^2 ô của một mảng các ô vuông kích thước $n \times n$ các số $1, 2, .., n$ sao cho mỗi số điền vào chỉ xuất hiện đúng một lần ở mỗi hàng và mỗi cột chứa số đó. Nói cách khác, mỗi hàng và mỗi cột này là một hoán vị của tập $\{1, 2, ..., n\}$. Khi đó ta nói đó là hình vuông Latin bậc n.

1	2	3	4
2	1	4	3
4	3	1	2
3	4	2	1

Bảng 4.1: Hình vuông Latin bậc 4.

Ta xét bài toán: Xác định số hình vuông latin $L(n)$ bậc n. Xét một vài ví dụ nhỏ :

$n = 1$: | 1 | : $L(1) = 1$

$n = 2$: | 1 | 2 | / | 2 | 1 | | | 2 | 1 | / | 1 | 2 | : $L(2) = 2$

$n = 3$:

1	2	3
2	3	1
3	1	2

: $L(3) = 12$

$$L(1) = 1, L(2) = 2, L(3) = 12, L(4) = 576, L(5) = 161280$$

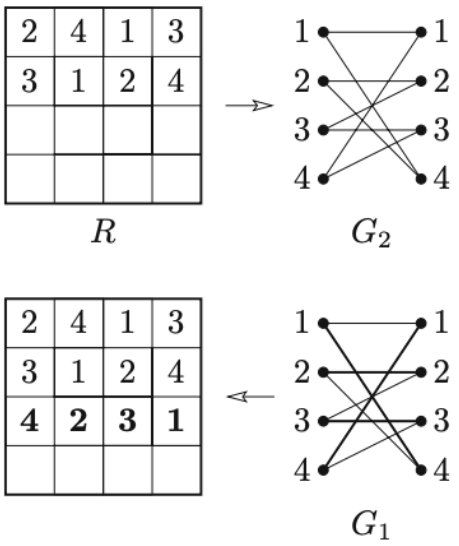
Ta thấy khi n lớn, thì số lượng hình vuông Latin bậc n tăng rất nhanh. Liệu ta có thể tìm được một chặn trên cho giá trị này ? Kết quả của định lý Brégman sẽ giải quyết được vấn đề này.

4.2 Mô hình bài toán

Xét một hình vuông $n \times n$ và điền vào đó các số $1, 2, .., n$ sao cho đó là một hình vuông Latin. Có $n!$ cách điền vào hàng đầu tiên (hoán vị của n phần tử). Giả sử $n - k$ hàng đầu tiên đã được điền và cho ta một hình chữ nhật Latin $(n - k) \times n$ R. Vậy có bao nhiêu cách để điền vào hàng tiếp theo? Xét một đồ thị hai phía $G_k = (U \cup V, E)$, trong đó U là tập các phần tử $\{1, 2, .., n\}$ và V là tập các vị trí của cột, sao cho:

$$ij \in E :\Leftrightarrow i \text{ không xuất hiện trong cột thứ } j \text{ của } R.$$

Nói cách khác, một cách điền phù hợp trong hàng tiếp theo tương ứng với một cặp ghép hoàn hảo (*perfect matching*) trong đồ thị hai phía $G_k = (U \cup V, E)$. Bây giờ, mọi thành phần $i \in U$ xuất hiện $n - k$ lần trong R nên nó sẽ xuất hiện trong k cột còn lại của hàng tiếp theo. Do đó, i có bậc k trong G_k và tương tự $d(j) = k$ với $j \in V$.



Đặt M_k là ma trận 0/1 tương ứng với G_k, khi đó:

$$\text{per } M_k = \text{số cách điền phù hợp của hàng } n-k+1.$$

Mọi hàng và cột trong M_k có tổng bằng k; ta kí hiệu tập các ma trận 0/1 bởi $\mathcal{M}(n,k)$. Giá trị *per* M_k phụ thuộc vào việc sắp xếp ma trận R, nhưng nếu ta có một cận trên và cận dưới hợp lý cho các ma trận trong $\mathcal{M}(n,k)$, khi ta lấy tích trên các giá trị của k, ta thu được một cận trên/cận dưới cho $L(n)$.
Sử dụng định lý Brégman với $d_1 = d_2 = ... = d_n = k$ ta có ngay

$$\text{per } M \leq k!^{\frac{n}{k}} \text{ với mọi } M \in \mathcal{M}(n,k)$$

Bây giờ là giá trị cho cận dưới: Giả sử $k < n$ và đặt L là hình chữ nhật Latin $k \times n$ (k hàng trên đã được điền) trên $\{1,2,...,n\}$. Để tính số cách điền các số vào các hàng tiếp theo $(k+1)$ để tạo thành hình chữ nhật Latin $(k+1) \times n$, đặt $M = (m_{ij})$ là một ma trận sao cho $m_{ij} = 1$ nếu i không xuất hiện ở cột thứ j và bằng 0 nếu ngược lại. Khi đó giá trị per (M) sẽ đếm số cách có thể điền vào hàng thứ $k+1$. Tổng theo hàng và cột của ma trận $M = n-k$ nên $\frac{1}{n-k}B$ là ma trận *ngẫu nhiên kép*. Theo định lý (1.1) về vĩnh thức: *per* $(M) \geq (n-k)^n \frac{n!}{n^n}$ nên ta có:

$$L(n) \geq n! \prod_{k=1}^{n-1} \{(n-k)^n \frac{n!}{n^n}\} = \frac{(n!)^{2n}}{n^{n^2}}.$$

Tổng quát, ta đã chứng minh được kết quả đáng chủ ý sau:

Định lý . *Giới hạn về số hình vuông Latin bậc n L(n) là*

$$\frac{n!^{2n}}{n^{n^2}} \leq L(n) \leq \prod_{k=1}^{n} k!^{\frac{n}{k}}.$$

Sử dụng tính xấp xỉ của $n!$:

$$\Big(\frac{n}{e}\Big)^n < n! < en\Big(\frac{n}{e}\Big)^n, \tag{4.1}$$

Ta có thể rút ra công thức về tiệm cận sau đây.

Hệ quả . *Theo giới hạn, số hình vuông Latin bậc n $L(n)$ thoả mãn*

$$\lim_{n\to\infty} \frac{L(n)^{\frac{1}{n^2}}}{n} = \frac{1}{e^2}$$

Chứng minh. Từ cận dưới của $L(n)$ ta có

$$L(n) \geq \frac{n!^{2n}}{n^{n^2}} > \frac{\Big(\frac{n}{e}\Big)^{2n^2}}{n^{n^2}} = \Big(\frac{n}{e^2}\Big)^{n^2}$$

nên $\frac{L(n)^{\frac{1}{n^2}}}{n} > \frac{1}{e^2}$ và do đó $\lim\limits_{n\to\infty} \frac{L(n)^{\frac{1}{n^2}}}{n} \geq \frac{1}{e^2}$
Để tìm cận trên, ta sẽ chỉ ra với mọi $\epsilon > 0$,

$$\frac{L(n)^{\frac{1}{n^2}}}{n} < \frac{1}{e^2}(1+\epsilon)$$

đúng khi n đủ lớn. Để thuận tiện ta đặt $\mathcal{L}(n) = L(n)^{\frac{1}{n^2}}$. Sử dụng (4.1) , ta có

$$\begin{aligned}
\log \mathcal{L}(n) &\leq \tfrac{1}{n} \log \prod_{k=1}^{n} (k!)^{\frac{1}{k}} = \frac{1}{n}\sum_{k=1}^{n} \frac{1}{k} \log k! \\
&< \tfrac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \log \Big(ek\Big(\frac{k}{e}\Big)^k\Big) \\
&= \tfrac{1}{n} \sum_{k=1}^{n} \frac{1}{k}(1 + \log k + k\log k - k) \\
&= \tfrac{1}{n} \Big[\sum_{k=1}^{n} \frac{1}{k} + \sum_{k=1}^{n} \frac{\log k}{k} + \sum_{k=1}^{n} \log k - n\Big].
\end{aligned}$$

Trong đó tổng đầu tiên là số *Harmonic* H_n: $H_n < \log n + 1$. Tổng thứ 3 được đánh giá thông qua (4.1), với:

$$\sum_{k=1}^{n} \log k < \log\Big(en\Big(\frac{n}{e}\Big)^n\Big) = 1 + (n+1)\log n - n \leq (n+2)\log n - n \tag{4.2}$$

Đối với tổng thứ hai, vì $\frac{\log x}{x}$ luôn dương với mọi $x > 1$ và làm hàm đơn điệu giảm với $x > e$, ta có:

$$\int_1^n \frac{\log x}{x} \geq \sum_{k=4}^{n} \int_{k-1}^{k} \frac{\log x}{x} dx \geq \sum_{k=4}^{n} \int_{k-1}^{k} \frac{\log k}{k} dx = \sum_{k=4}^{n} \frac{\log k}{k},$$

nên

$$\sum_{k=1}^{n} \frac{\log k}{k} \leq 2 + \left[\frac{1}{2}(\log x)^2\right] |_1 \, n = 2 + \frac{1}{2}(\log n)^2. \tag{4.3}$$

Từ (4.2) và (4.3) ta có:

$$\log \mathcal{L}(n) < \frac{3 \log n}{n} + \frac{3}{n} + \frac{(\log n)^2}{2n} + \log n - 2,$$

trong đó 3 số hạng đầu tiến dần về 0 khi n đủ lớn. Do đó, với mọi $\delta > 0$:

$$\log \mathcal{L}(n) \leq \delta + \log n - 2.$$

nên $L(n)^{\frac{1}{n^2}} \leq \frac{n}{e^2} e^{\delta}$ khi n đủ lớn. Ta có điều phải chứng minh. □