

Định lý Van der Waerden

Trương Phước Nhân, 08/05/2018

Ma trận thực $A = (a_{ij})_{n \times n}$ được gọi là song ngẫu nhiên nếu

$$a_{ij} \geq 0 \text{ với mọi } i, j \in \{1, 2, \dots, n\},$$

$$\sum_{j=1}^n a_{ij} = 1, \quad i = 1, 2, \dots, n,$$

$$\sum_{i=1}^n a_{ij} = 1, \quad i = 1, 2, \dots, n.$$

Bài toán. (Van der Waerden) Giả sử A là một ma trận song ngẫu nhiên cấp n . Khi đó, $\text{per}(A) \geq \frac{n!}{n^n}$.

Dấu “=” xảy ra khi và chỉ khi $A = \frac{1}{n} J$ với J là ma trận vuông cấp n với tất cả các phần tử đều bằng 1.

Đầu tiên, ta nhắc lại một số khái niệm và công cụ toán học cần thiết cho quá trình lập luận như sau:

Hàm $f(x_1, \dots, x_n)$ được gọi là *song tuyến tính* nếu $f(x_1, \dots, x_n) = \sum_{i,j} f_{ij} x_i x_j$. Thông thường ta biểu diễn

dạng song tuyến tính bởi một ma trận cấp $n \times n$ với phần tử ở vị trí (i, j) là f_{ij} . Do hệ số của $x_i x_j$ với $i \neq j$ trong khai triển của $f(x_1, \dots, x_n)$ là $f_{ij} + f_{ji}$ nên ta có thể giả sử $f_{ij} = f_{ji}$ và cùng bằng một nửa hệ số của $x_i x_j$; khi đó ma trận $F = (f_{ij})_{i,j=1,n}$ đối xứng.

Tiếp theo ta sẽ khảo sát xem dạng song tuyến tính sẽ biến đổi như thế nào khi ta thực hiện phép biến đổi tuyến tính lên các biến, có nghĩa là ta sẽ thay x_i bởi $\sum c_{ij} y_j$ với $i = 1, \dots, n$, điều đó sẽ biến đổi $f(x_1, \dots, x_n)$ thành dạng song tuyến tính $g(y_1, \dots, y_n)$. Nếu ma trận của phép biến đổi không suy biến thì ta có thể tìm được ma trận nghịch đảo của nó và do đó ta có thể biểu diễn các y_j theo các x_i và khi đó hai dạng song tuyến tính là tương đương.

Ta sẽ chứng minh rằng mọi dạng song tuyến tính đều tương đương với một dạng song tuyến tính “đơn giản”. Ý tưởng này là một sự mở rộng cho trường hợp tổng quát của phép “bình phương đủ” dùng cho các phương trình bậc hai.

Bổ đề 1. Cho dạng song tuyến tính $f(x_1, \dots, x_n) = \sum_{i,j} f_{ij} x_i x_j$ và vector $(a_1, \dots, a_n)$ sao cho $a_1 \neq 0$,

$$f(a_1, \dots, a_n) = c \neq 0.$$

Khi đó phép biến đổi không suy biến xác định bởi

$$x_1 = a_1 \left(y_1 - \sum_{j \geq 2} y_j \sum_r \frac{f_{jr} a_r}{c} \right),$$

$$x_i = a_i \left(y_1 - \sum_{j \geq 2} y_j \sum_r \frac{f_{jr} a_r}{c} \right) + y_i \text{ với } (i \geq 2), \quad (1)$$

$$\text{và } y_1 = \sum_{i,j} \frac{f_{ij} a_i x_j}{c},$$

$$y_i = x_i - \frac{x_1 a_i}{a_1} \text{ với } (i \geq 2), \quad (2)$$

biến $f(x_1, \dots, x_n)$ thành $cy_1^2 + g(y_2, \dots, y_n)$, trong đó g là một dạng song tuyến tính theo $(n-1)$ biến.

Chứng minh. Dễ dàng chứng minh được rằng (1) và (2) là nghịch đảo của nhau. Đặt c_{ij} là hệ số của y_j trong x_i . Hệ số của y_1^2 trong $f(x_1, \dots, x_n)$ bằng $\sum_{i,j} f_{ij} c_{i1} c_{j1} = \sum_{i,j} f_{ij} a_i a_j = c$ và hệ số của $y_1 y_k$, với $k \geq 2$, là

$$\sum_{i,j} f_{ij} (c_{i1} c_{jk} + c_{j1} c_{ik}) = \sum_{i,j} f_{ij} a_i \left(\delta_{jk} - \frac{a_j}{c} \sum_r f_{kr} a_r \right) + \sum_{i,j} f_{ij} a_j \left(\delta_{ik} - \frac{a_i}{c} \sum_r f_{kr} a_r \right) \text{ triệt tiêu thành } 0, \text{ trong đó } \delta_{ij} \text{ là}$$

kí hiệu Kronecker.

Bổ đề 2. Mọi dạng song tuyến tính $f(x_1, \dots, x_n)$ đều tương đương với một dạng song tuyến tính “đơn giản” $g(y_1, \dots, y_n) = y_1^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2$ với $0 \leq p \leq r \leq n$.

Chứng minh. Nếu $f(a_1, \dots, a_n)$ triệt tiêu với mọi $(a_1, \dots, a_n)$ thì $f_{ij} = 0$ với mọi i, j và ta thu được kết luận với $p = r = 0$. Nếu $f(a_1, \dots, a_n)$ khác không tại ít nhất một vector $(a_1, \dots, a_n)$ thì $a_i \neq 0$; bằng cách sắp xếp lại số thứ tự các biến (nếu cần) ta có thể giả sử $a_1 \neq 0$. Bây giờ bằng cách áp dụng phương pháp được sử dụng trong Bổ đề 1 nhưng ta thay y_1 bởi $z_1 = y_1 |c|^{-\frac{1}{2}}$, ta thu được $f(x_1, \dots, x_n) = \pm z_1^2 + g(y_2, \dots, y_n)$. Kết luận của bổ đề được suy ra đơn giản thông qua một phép quy nạp.

Bổ đề 3. Các số p và r nêu trong Bổ đề 2 được xác định duy nhất, theo nghĩa là, nếu ta có hai dạng song tuyến tính tương đương nhau $y_1^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 = z_1^2 + \dots + z_q^2 - z_{q+1}^2 - \dots - z_s^2$ thì $p = q$ và $r = s$.

Chứng minh. Đặt $y_i = \sum_j c_{ij} z_j$ và giả sử rằng $p < q$. Khi đó ta có thể tìm được các giá trị $z_1^*, \dots, z_q^*$ không đồng thời bằng không sao cho $\sum_{1 \leq j \leq q} c_{ij} z_j^* = 0$ với $1 \leq i \leq p$, bởi vì đây là một hệ thuần nhất gồm p phương trình tuyến tính theo $q > p$ ẩn số z_j^* . Đặt $z_{q+1}^* = \dots = z_s^* = 0$ và $y_i^* = \sum_{1 \leq j \leq s} c_{ij} z_j^*$ với mọi i .

Khi đó $-y_{p+1}^{*2} - \dots - y_r^{*2} = z_1^{*2} + \dots + z_q^{*2}$, điều này chỉ có thể xảy ra khi $z_1^* = \dots = z_q^* = 0$, mâu thuẫn với cách chọn các giá trị $z_1^*, \dots, z_q^*$. Do đó $p = q$. Bây giờ nếu $r < s$, bằng phương pháp lập luận tương tự ta có thể tìm được các giá trị $z_{q+1}^*, \dots, z_s^*$ không đồng thời bằng không sao cho $\sum_{q+1 \leq j \leq s} c_{ij} z_j^* = 0$ với $p+1 \leq i \leq r$, và đặt $z_1^* = \dots = z_q^* = 0$ ta thu được $y_1^{*2} + \dots + y_p^{*2} = -z_{q+1}^{*2} - \dots - z_s^{*2}$, mâu thuẫn với cách chọn.

Từ Bổ đề 3 ta suy ra rằng các số p và r là các đại lượng bất biến xác định dạng song tuyến tính f , thông thường ta kí hiệu là $p(f)$ và $r(f)$.

Bổ đề 4. Cho $f_\theta(x_1, \dots, x_n)$ là dạng song tuyến tính xác định bởi

$$f_\theta(x_1, \dots, x_n) = (1 - \theta) f_0(x_1, \dots, x_n) + \theta f_1(x_1, \dots, x_n)$$

biến đổi dạng song tuyến tính f_0 thành f_1 khi θ thay đổi từ 0 đến 1.

Khi đó, nếu $r(f_\theta) = n$ với $0 \leq \theta \leq 1$ thì $p(f_0) = p(f_1)$.

Chứng minh. Để đơn giản cho quá trình trình bày ta chỉ phác thảo sơ bộ phép chứng minh của Bổ đề 4. Dễ dàng nhận thấy rằng $p(f_\theta)$ là số các giá trị riêng dương của ma trận đối xứng s_θ liên kết với dạng song tuyến tính f_θ . Bây giờ $r(f_\theta) = n$ có nghĩa là ma trận s_θ không có giá trị riêng bằng 0. Khẳng định của Bổ đề 4 được suy ra từ sự kiện các nghiệm của một đa thức là các hàm liên tục theo các hệ số và do đó các giá trị riêng của s_θ cũng sẽ là các hàm liên tục theo θ , lưu ý các hệ số của đa thức đặc trưng của s_θ là các đa thức theo θ .

Tiếp theo, ta xét dạng song tuyến tính liên kết với permanent. Hai bổ đề sau được suy ra đơn giản bằng phép quy nạp theo n .

Bổ đề 5. Cho $a_1, \dots, a_{n-1}$ là các vector với các thành phần không âm sao cho vector a_i có ít nhất $n+1-i$ thành phần dương và giả sử $b = (b_1, \dots, b_n)$ là vector với các thành phần thực sao cho $\text{per}(a_1, \dots, a_{n-1}, b) = 0$. Khi đó $\text{per}(a_1, \dots, a_{n-2}, b, b) \leq 0$.

Dấu “=” xảy ra khi và chỉ khi $b_1 = b_2 = \dots = b_n = 0$, trong đó $\text{per}(c_1, \dots, c_n)$ là permanent của ma trận có các hàng là các vector $c_1, \dots, c_n$.

Bổ đề 6. Cho $a_1, \dots, a_{n-2}$ là xác định như trong Bổ đề 5 và f là dạng song tuyến tính

$$f(x_1, \dots, x_n) = \text{per}(a_1, \dots, a_{n-2}, x, x), \quad (3)$$

trong đó $x = (x_1, \dots, x_n)$. Khi đó $r(f) = n$ và $p(f) = 1$.

Chứng minh.

Đầu tiên xét $n = 2$, với $a_1 = (a_{11}, a_{12})$ sao cho $a_{11} > 0$, $a_{12} > 0$, và $a_{11}b_2 + a_{12}b_1 = 0$.

Khi đó $\text{per}(b, b) = 2b_1b_2 \leq 0$. Dấu “=” xảy ra khi và chỉ khi $b_1 = b_2 = 0$.

Hơn nữa, $\text{per}(x, x) = 2x_1x_2 = \left(\frac{x_1+x_2}{\sqrt{2}}\right)^2 - \left(\frac{x_1-x_2}{\sqrt{2}}\right)^2$ nên $r(f) = 2$ và $p(f) = 1$.

Giả sử $n \geq 3$ và hai bổ đề đã được kiểm chứng với $n-1$.

Trong dạng song tuyến tính (3) giá trị của f_{ij} là permanent của ma trận cấp $(n-2) \times (n-2)$ thu được bằng cách xóa đi hàng thứ i và cột thứ j từ ma trận $(a_1, \dots, a_{n-2})$ nếu $i \neq j$ và đồng thời $f_{ii} = 0$. Nếu $r(f) < n$, ma trận tương ứng của dạng song tuyến tính f là suy biến, nên ta tìm được một vector $(c_1, \dots, c_n) \neq 0$ sao cho $\sum_j f_{ij}c_j = 0$ với mọi i . Điều này tương đương với sự kiện $\text{per}(a_1, \dots, a_{n-2}, c, x) = 0$ với mọi x , lưu ý đến ý nghĩa tổ hợp của f_{ij} ; trong trường hợp riêng $\text{per}(a_1, \dots, a_{n-2}, c, c) = 0$. Hơn nữa, $\text{per}_j(a_1, \dots, a_{n-2}, c) = 0$ với mọi j , trong đó per_j là permanent thu được bằng cách xóa đi cột thứ j . Bằng quy nạp, $\text{per}_j(a_1, \dots, a_{n-3}, c, c) \leq 0$ với mọi j , dấu “=” xảy ra khi và chỉ khi $c_k = 0$ với mọi $k \neq j$, với mọi chỉ số j . Khi đó

$$0 = \text{per}(a_1, \dots, a_{n-2}, c, c) = \sum_j a_{n-2,j} \text{per}_j(a_1, \dots, a_{n-3}, c, c) \leq 0.$$

Do đó $\text{per}_j(a_1, \dots, a_{n-3}, c, c) = 0$ khi $a_{n-2,j} > 0$. Điều này xảy ra đối với ít nhất hai giá trị của j (trên thực tế thì nó xảy ra đối với ít nhất là ba giá trị của j); do đó $c_1 = \dots = c_n = 0$, mâu thuẫn với cách chọn c .

Như vậy chúng ta đã chứng minh được nửa về đầu của Bổ đề 6, nghĩa là $r(f) = n$. Đối với nửa sau, xét $f_0 = \text{per}(a_1, a_2, \dots, a_{n-2}, x, x)$ và $f_1 = \text{per}(\bar{1}, a_2, \dots, a_{n-2}, x, x)$ trong đó $\bar{1} = (1, \dots, 1)$. Đặt $f_\theta = (1-\theta)f_0 + \theta f_1$. Từ Bổ đề 4, nếu $r(f_\theta) = n$, ta suy ra $p(f_0) = p(f_1)$. Theo cách này ta có thể thay $a_1, a_2, \dots, a_{n-2}$ bởi vector $(1, \dots, 1)$ và khi đó ta chỉ cần tính $p(f)$ cho dạng song tuyến tính $f = \text{per}(\bar{1}, \bar{1}, \dots, \bar{1}, x, x)$.

Khi đó $f_{ij} = (n-2)!$ nếu $i \neq j$ và $f_{ii} = 0$, nên ta chỉ cần xét

$$\frac{1}{(n-2)!} f = \sum_{i \neq j} x_i x_j = (n-1)s^2 - (x_2 - s)^2 - \dots - (x_n - s)^2,$$

trong đó $s = x_1 + \frac{1}{2}(x_2 + \dots + x_n)$. Như vậy $p(f) = 1$ và như vậy là ta đã chứng minh được hoàn tất Bổ đề 6 với mọi n .

Bây giờ ta quay lại với Bổ đề 5, từ giả thiết kéo theo rằng $\text{per}(a_1, \dots, a_{n-1}, a_{n-1}) > 0$.

Do đó $f(a_{(n-1),1}, a_{(n-1),2}, \dots, a_{(n-1),n}) = c > 0$. Không mất tính tổng quát, bằng cách hoán vị các cột, ta có thể giả sử $a_{(n-1),1} > 0$. Bằng cách áp dụng cách xây dựng được đề cập trong Bổ đề 1, ta thu được

$$f(x_1, \dots, x_n) = cy_1^2 + g(y_2, \dots, y_n).$$

Bằng cách áp dụng Bổ đề 6, $g(y_2, \dots, y_n) \leq 0$ với mọi $(y_2, \dots, y_n)$, và nó đồng nhất bằng không khi và chỉ khi $y_2 = \dots = y_n = 0$.

Sử dụng (2),

$$y_1 = \sum_{i,j} \frac{f_{ij} a_{(n-1),i} x_j}{c} = \frac{\text{per}(a_1, \dots, a_{n-1}, x)}{c}.$$

Từ giả thiết của Bổ đề 5 ta suy ra $f(b_1, \dots, b_n) \leq 0$. Sau cùng $f(b_1, \dots, b_n) = 0$ kéo theo $y_1 = \dots = 0$ khi $(x_1, \dots, x_n) = (b_1, \dots, b_n)$; nên $b_1 = b_2 = \dots = b_n = 0$.

Sau đây ta sẽ trình bày một số kết quả chính yếu dùng cho phép chứng minh của định lý Van der Waerden

Kết quả 1. Cho $a_1, \dots, a_{n-1}$ là các vector với các thành phần không âm sao cho a_i có ít nhất $(n+1-i)$ thành phần dương và a_n là một vector với các thành phần thực tùy ý.

$$\text{Khi đó } \text{per}(a_1, \dots, a_{n-1}, a_n)^2 \geq \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \text{per}(a_1, \dots, a_{n-2}, a_n, a_n)$$

Dấu “=” xảy ra khi và chỉ khi $a_n = \lambda a_{n-1}$ với một số thực λ nào đó.

Chứng minh. Đặt $\text{per}(a_1, \dots, a_{n-1}, a_n) = \lambda \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1})$, khi đó λ được xác định bởi vì $\text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) > 0$. Nếu ta đặt $b = a_n - \lambda a_{n-1}$ thì ta có $\text{per}(a_1, \dots, a_{n-1}, b) = 0$, bởi vì permanent là một hàm tuyến tính theo mỗi hàng của nó. Bằng cách áp dụng Bổ đề 5, ta suy ra

$$\begin{aligned} 0 &\geq \text{per}(a_1, \dots, a_{n-2}, b, b) \\ &= \text{per}(a_1, \dots, a_{n-2}, b, a_n) - \lambda \text{per}(a_1, \dots, a_{n-2}, b, a_{n-1}) \\ &= \text{per}(a_1, \dots, a_{n-2}, a_n, a_n) - 2\lambda \text{per}(a_1, \dots, a_{n-2}, b, a_{n-1}) + \lambda^2 \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \\ &= \text{per}(a_1, \dots, a_{n-2}, a_n, a_n) - \lambda^2 \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \\ &= \text{per}(a_1, \dots, a_{n-2}, a_n, a_n) - \frac{\text{per}(a_1, \dots, a_n)^2}{\text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1})} \end{aligned}$$

Bằng cách áp dụng Bổ đề 5, ta suy ra

$$\text{per}(a_1, \dots, a_{n-1}, a_n)^2 \geq \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \text{per}(a_1, \dots, a_{n-2}, a_n, a_n),$$

dấu “=” xảy ra khi và chỉ khi $b = 0$, hay $a_n = \lambda a_{n-1}$.

Từ Kết quả 1 ta suy ra hệ quả sau

Hệ quả. Nếu $a_1, \dots, a_{n-1}$ là các vector có các thành phần không âm và a_n là một vector bất kì thì ta có

$$\text{per}(a_1, \dots, a_{n-1}, a_n)^2 \geq \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \text{per}(a_1, \dots, a_{n-2}, a_n, a_n)$$

Chứng minh. Bằng cách áp dụng Kết quả 1 cho các vector $a_i + (\varepsilon, \dots, \varepsilon)$ và cho ε tiến dần tới không ta thu được điều phải chứng minh.

Cho trước ma trận song ngẫu nhiên $A = (a_{ij})_{i,j=1,\dots,n}$, vector $x = (x_1, \dots, x_n)$ được là **cân bằng** nếu

$Ax^T = x^T$, có nghĩa là $\sum_j a_{ij} x_j = x_i$ với mọi i . Ma trận A được gọi là **decomposable** nếu tập $\{1, \dots, n\}$ có thể

phân hoạch thành hai tập con S và T khác rỗng sao cho $a_{ij} = 0$ với $i \in S$ và $j \in T$. Sau đây ta sẽ trình bày một số khảo sát đơn giản về mối liên hệ giữa hai khái niệm vừa trình bày.

Bổ đề 7. Nếu $x = (x_1, \dots, x_n)$ là một vector cân bằng của ma trận A với một số thành phần không bằng nhau thì A là decomposable.

Chứng minh. Nếu $x = (x_1, \dots, x_n)$ là một vector cân bằng thì $(x_1 + c, \dots, x_n + c)$ cũng là một vector cân bằng; từ giả thiết không làm mất tính tổng quát ta có thể giả sử rằng tất cả các thành phần của x đều không âm và một số thành phần, không phải tất cả, bằng không. Đặt $S = \{i \mid x_i = 0\}$ và $T = \{i \mid x_i > 0\}$.

Khi đó hệ thức $\sum_j a_{ij} x_j = x_i$ kéo theo rằng $a_{ij} = 0$ với $i \in S$ và $j \in T$.

Nhận thấy rằng chứng minh của Bổ đề 7 vừa trình bày ở trên chỉ cần đến tính “ngẫu nhiên” $\sum_j a_{ij} = 1$ với

$1 \leq i \leq n$. Ma trận song ngẫu nhiên thỏa mãn một điều kiện còn mạnh hơn rất nhiều.

Bổ đề 8. Cho A là một ma trận song ngẫu nhiên và nếu S và T lập thành một decomposable của A thì $a_{ij} = 0$ ngoại trừ trường hợp $i, j \in S$ hoặc $i, j \in T$.

Chứng minh. Ta có

$$\sum_{i,j \in S} a_{ij} = \sum_{i \in S} \sum_{1 \leq j \leq n} a_{ij} = \sum_{i \in S} 1 = \sum_{j \in S} 1 = \sum_{j \in S} \sum_{1 \leq i \leq n} a_{ij} = \sum_{i,j \in S} a_{ij} + \sum_{i \in T, j \in S} a_{ij},$$

nên $a_{ij} = 0$ khi $i \in T$ hoặc $j \in S$.

Một ma trận $A = (a_{ij})$ cấp $n \times n$ được gọi là **minimal** nếu nó là song ngẫu nhiên và nó có permanent nhỏ nhất trong số tất cả các ma trận song ngẫu nhiên cấp $n \times n$. Đầu tiên, ta nhận thấy rằng luôn có ít nhất một

ma trận minimal có cấp n , bởi vì permanent là một hàm liên tục theo các phần tử của ma trận và tập các ma trận song ngẫu nhiên là một tập con đóng, bị chặn của không gian n^2 - chiều. Bây giờ ta để ý rằng nội dung của định lý Van der Waerden chỉ ra rằng chỉ có duy nhất một ma trận minimal có cấp n , đó chính là ma trận có $a_{ij} = \frac{1}{n}$ với mọi i, j . Tiếp theo, để ý rằng $\text{per} A$ luôn dương nên giá trị nhỏ nhất không thể bằng không.

Đặt A_{ij} là ma trận cấp $(n-1) \times (n-1)$ thu được từ ma trận A bằng cách xóa đi hàng thứ i và cột thứ j .

Theo một kết quả quen thuộc về permanent ta có $\text{per}(A) = \sum_j a_{ij} \text{per}(A_{ij}) = \sum_i a_{ij} \text{per}(A_{ij})$.

Từ hệ thức trên ta dễ dàng thu được $\text{per}(A + \varepsilon B) = \text{per}(A) + \varepsilon \sum_{i,j} b_{ij} \text{per}(A_{ij}) + 0(\varepsilon^2)$.

Cho A là một ma trận song ngẫu nhiên ta nói B là **valid modification** của A nếu tổng các hàng và tổng các cột của B đều bằng không và nếu $a_{ij} = 0$ thì $b_{ij} \geq 0$; khi đó $A + \varepsilon B$ là ma trận song ngẫu nhiên với mọi $\varepsilon > 0$ đủ nhỏ.

Bổ đề 9. Nếu A là một ma trận minimal và B là ma trận valid modification của A thì

$$\sum_{i,j} b_{ij} \text{per}(A_{ij}) \geq 0$$

Chứng minh. Khẳng định trên được suy ra từ hệ thức $\text{per}(A + \varepsilon B) = \text{per}(A) + \varepsilon \sum_{i,j} b_{ij} \text{per}(A_{ij}) + 0(\varepsilon^2)$ và định nghĩa của ma trận minimal.

Bổ đề 10. Ma trận minimal không thể là decomposable.

Chứng minh. Giả sử A là ma trận song ngẫu nhiên và $a_{ij} > 0$ chỉ khi $i, j \in S$ hoặc $i, j \in T$, trong đó S và T là các tập khác rỗng và có hợp bằng với tập $\{1, 2, \dots, n\}$. Ta biết rằng $\text{per}(A) > 0$ nên tồn tại một hoán vị $\pi(1), \dots, \pi(n)$ sao cho $a_{i\pi(i)} > 0$ với mọi i . Gọi s và t lần lượt là hai phần tử của S và T và đặt $B = (b_{ij})$ là ma trận với hầu hết tất cả các phần tử đều bằng không ngoại trừ $b_{s\pi(s)} = b_{t\pi(t)} = -1$ và $b_{s\pi(t)} = b_{t\pi(s)} = 1$.

Do $\pi(s) \in S$ và $\pi(t) \in T$ nên ma trận B là valid modification của A , nên

$$-\text{per}(A_{s\pi(s)}) - \text{per}(A_{t\pi(t)}) + \text{per}(A_{s\pi(t)}) + \text{per}(A_{t\pi(s)}) \geq 0, \text{ suy ra từ Bổ đề 9.}$$

Nhưng điều này là không thể xảy ra bởi vì $\text{per}(A_{s\pi(s)})$ và $\text{per}(A_{t\pi(t)})$ là dương còn $\text{per}(A_{s\pi(t)})$ và $\text{per}(A_{t\pi(s)})$ bằng không. Để đơn giản ta sẽ chỉ dẫn giải cho $\text{per}(A_{t\pi(s)})$ trường hợp còn lại làm tương tự. Do ma trận $A_{t\pi(s)}$ có k hàng tương ứng với S , ở đây ta giả sử $|S| = k$, trong khi tất cả các phần tử dương lại xuất hiện chỉ trong $(k-1)$ cột tương ứng với $S \setminus \{\pi(s)\}$.

Từ các lập luận trên ta suy ra rằng ma trận A không thể là decomposable.

Bổ đề 11. Nếu A là ma trận minimal thì $\text{per}(A_{ij}) > 0$ với mọi i, j .

Chứng minh. Nếu $\text{per}(A_{ij}) = 0$ thì từ ma trận A_{ij} ta tìm được tập S gồm $k > 0$ cột sao cho tất cả các phần tử dương của ma trận xuất hiện trong $(k-1)$ cột tương ứng với $S \setminus \{j\}$ theo định lý Hall.

Đặt $T = \{1, \dots, n\} \setminus S$; chú ý rằng T không rỗng, bởi vì $i \in T$. Bằng cách hoán vị các hàng và các cột của ma trận A ta thu được ma trận A' sao cho S chứa k hàng đầu còn T chứa $(n-k)$ vị trí còn lại sao cho tất cả các phần tử dương của k hàng đầu nằm trong k cột đầu và chúng chính là cột j ban đầu và $(k-1)$ cột dương của A_{ij} . Hiển nhiên A' cũng là một ma trận minimal. Mặt khác theo cách giải thích nêu trên thì A' là decomposable, mâu thuẫn với Bổ đề 10.

Kết quả 2. Nếu A là ma trận minimal và $a_{ij} > 0$ thì $\text{per}(A_{ij}) = \text{per}(A)$.

Chứng minh. Ý tưởng cơ bản của chứng minh này là cần sử dụng đến tập tất cả các ma trận B là valid modifications của A và cố gắng chỉ ra rằng $b_{ij} = 0$ khi $a_{ij} = 0$. Đầu tiên ta sẽ chỉ ra rằng tồn tại các hằng số $\lambda_1, \dots, \lambda_n, \mu_1, \mu_2, \dots, \mu_n$ sao cho $\text{per}(A_{ij}) = \lambda_i + \mu_j$ nếu $a_{ij} > 0$.

Không mất tính tổng quát, bằng cách hoán vị các cột ta có thể giả sử rằng $a_{ii} > 0$ với mọi i . Ta kí hiệu $i \rightarrow j$ nếu $a_{ij} > 0$; do đó $i \rightarrow i$. Do A không thể là decomposable nên tồn tại một “đường đi”

$$1 = j_0 \rightarrow j_1 \rightarrow \dots \rightarrow j_t = j$$

từ 1 đến j với mọi j . Ta nói rằng j ở khoảng cách s nếu đường đi ngắn nhất như vừa trình bày ở trên có độ dài là s . Với mọi $j > 1$ ở khoảng cách s , đặt $p(j)$ là chỉ số nhỏ nhất ở khoảng cách $(s-1)$ sao cho $p(j) \rightarrow j$. Khi đó đường đi $1 = j_0 \rightarrow j_1 \rightarrow \dots \rightarrow j_t = j$ được xác định duy nhất nếu ta ngầm hiểu $j_{k-1} = p(j_k)$ với $1 \leq k \leq s$. Các cung $\{p(j) \rightarrow j \mid j > 1\}$ có thể được xem như là một cây định hướng xuất phát từ điểm 1. Ta nói rằng i là tổ tiên của j (và ta kí hiệu là $i < j$) nếu ta có $i = p(j)$, hoặc $p(p(j))$ hoặc $i = p(\dots(p(j)))$, đồng thời ta cũng kí hiệu $i \preceq j$ khi muốn diễn đạt rằng $i = j$ hoặc $i < j$.

Bây giờ ta sẽ tiến hành xây dựng các hằng số $\lambda_1, \dots, \lambda_n, \mu_1, \mu_2, \dots, \mu_n$ như sau:

Đầu tiên, ta đặt $\lambda_1 = 0$ và $\mu_1 = \text{per}(A_{11})$. Với các chỉ số $j \geq 2$ được sắp xếp theo thứ tự tăng dần của khoảng cách từ 1 đến chúng và ta có thể xây dựng các giá trị μ_j được sao cho $\text{per}(A_{ij}) = \lambda_i + \mu_j$ luôn đúng trong đó $i = p(j)$, bởi vì giá trị của $\lambda_{p(j)}$ đã được xây dựng trước đó rồi; sau khi đã xây dựng μ_j ta sẽ tiến hành xây dựng tiếp λ_j sao cho hệ thức $\text{per}(A_{ij}) = \lambda_i + \mu_j$ vẫn được thỏa mãn với $i = j$.

Tiếp theo, ta sẽ kiểm lại hệ thức $\text{per}(A_{ij}) = \lambda_i + \mu_j$ với $a_{ij} > 0$ cho các bộ (h, k) với $a_{hk} > 0$, $h \neq k$ và $h \neq p(k)$. Để làm điều này ta xét ma trận B có hầu hết tất cả các phần tử bằng không ngoại trừ

$$\begin{aligned} b_{hk} &= 1, \\ b_{jj} &= (j < k) - (j \preceq h) \quad \text{với } 1 \leq j \leq n, \\ b_{p(j)j} &= (j \preceq h) - (j \preceq k) \quad \text{với } 1 \leq j \leq n, \end{aligned}$$

trong đó kí hiệu $(j < k)$ nhận hai giá trị 1 hoặc 0 phụ thuộc vào hệ thức $j < k$ là đúng hoặc sai các trường hợp còn lại thì tương tự.

Nhận thấy $\sum_i b_{ij} = (j = k) + ((j < k) - (j \preceq h) + (j \preceq h) - (j \preceq k)) = 0$ và

$\sum_j b_{ij} = (i = h) + ((i < k) - (i \preceq h)) + ((i < h) - (i < k)) = 0$; trong hệ thức sau ta đã sử dụng một sự thật rằng với i cố định luôn tồn tại một giá trị j sao cho $i = p(j)$ và $j < h$ nếu và chỉ nếu $i \preceq h$. Từ điều vừa trình bày ta nhận ra rằng B là valid modification của ma trận A và tương tự của $-B$ cũng như vậy. Bằng cách áp dụng Bổ đề 9 ta nhận thấy rằng $\sum_{i,j} b_{ij} \text{per}(A_{ij})$ bằng không. Do đó

$$\sum_{i,j} b_{ij} (\text{per}(A_{ij}) - \lambda_i - \mu_j) = \text{per}(A_{hk}) - \lambda_h - \mu_k,$$

bởi vì $\lambda_i \sum_j b_{ij} = \mu_j \sum_i b_{ij} = 0$ và bởi một phép quy nạp đơn giản, hệ thức $\text{per}(A_{ij}) = \lambda_i + \mu_j$ được thỏa mãn với mọi cặp (i, j) đã được kiểm chứng trước đó với $b_{ij} \neq 0$ chỉ ngoại trừ đối với cặp (h, k) cho trước thì khẳng định này chưa thật rõ ràng. Từ hệ thức vừa dẫn trên ta nhận thấy hệ thức $\text{per}(A_{ij}) = \lambda_i + \mu_j$ được thỏa mãn trong trường hợp tổng quát.

Bây giờ ta sẽ hoàn chỉnh nốt phần việc mà ta đã đề ra ngay từ đây. Đầu tiên, ta có

$a_{ij} \text{per}(A_{ij}) = a_{ij} (\lambda_i + \mu_j)$ với mọi i, j . Bằng cách sử dụng hệ thức $\text{per}(A) = \sum_j a_{ij} \text{per}(A_{ij})$ ta thu được

$$\text{per}(A) = \lambda_i + \sum_j a_{ij} \mu_j = \mu_i + \sum_j a_{ij} \lambda_j \quad \text{với mọi } i, j.$$

Nếu ta đặt $\lambda = (\lambda_1, \dots, \lambda_n)^T$, $\mu = (\mu_1, \dots, \mu_n)^T$ và $e = (1, \dots, 1)^T$, khi đó các kết quả trên được biểu diễn lại dưới dạng ma trận như sau $\lambda + A\mu = \mu + A^T \lambda = \text{per}(A)e$. Do $Ae = A^T e = e$ nên ta có $A^T \lambda + A^T A\mu = \text{per}(A)e$ và $A\mu + AA^T \lambda = \text{per}(A)e$, tổng hợp các điều trên lại ta thu được $\mu = A^T A\mu$ và $\lambda = AA^T \lambda$.

Bây giờ, do A không phải là decomposable và có các phần tử dương nằm trên đường chéo chính nên các ma trận $A^T A$ và AA^T , lưu ý rằng chúng cũng là các ma trận song ngẫu nhiên, thậm chí cũng là các ma trận không decomposable. Bằng cách áp dụng Bổ đề 7 ta suy ra $\lambda_1 = \dots = \lambda_n$ và $\mu_1 = \dots = \mu_n$. Nếu ban đầu ta chọn $\lambda_1 = 0$ thì $\lambda_1 = \dots = \lambda_n = 0$ và từ cách xây dựng ta nhận thấy $\mu = \text{per}(A)e$ và do đó $\mu_1 = \dots = \mu_n = \text{per}(A)$. Điều này chỉ khẳng định cần chứng minh.

Bổ đề 12. Nếu ma trận A là minimal thì $\text{per}(A_{ij}) \geq \text{per}(A)$ với mọi i, j .

Chứng minh. Từ Kết quả 2 ta nhận thấy rằng chỉ cần xét các cặp (i, j) sao cho $a_{ij} = 0$. Không mất tính tổng quát ta có thể giả sử $i = j = 1$ và $a_{11} > 0$. Bằng cách áp dụng Bổ đề 11 $\text{per}(A_{11}) > 0$ và bằng cách hoán vị các cột ta có thể giả sử rằng $a_{jj} > 0$ và đặt $B = I - A$; tất cả các phần tử của B tương ứng với $a_{ij} = 0$ đều không âm, tổng các hàng và các cột đều bằng không nên B là ma trận valid modification của A .

Áp dụng Bổ đề 9 ta thu được

$$\begin{aligned} 0 &\leq \sum_{i,j} b_{ij} \text{per}(A_{ij}) = \sum_j \text{per}(A_{jj}) - \sum_{i,j} a_{ij} \text{per}(A_{ij}) \\ &= \text{per}(A_{11}) + (n-1) \text{per}(A) - n \text{per}(A) = \text{per}(A_{11}) - \text{per}(A), \end{aligned}$$

do $\text{per}(A_{jj}) = \text{per}(A)$ với $j > 1$ theo Kết quả 2 và sử dụng đến hệ thức $\text{per}(A) = \sum_j a_{jj} \text{per}(A_{jj})$.

Sau cùng, để kết thúc bài viết ta sẽ chứng minh hoàn chỉnh định lý Van der Waerden đã nêu ở đầu đề mục

Bổ đề 13. Nếu A là ma trận minimal thì $\text{per}(A_{ij}) = \text{per}(A)$ với mọi i, j .

Chứng minh.

Không mất tính tổng quát ta giả sử rằng $i = j = n$ và giả sử thêm rằng $\text{per}(A_{nn}) > \text{per}(A)$, $a_{(n-1),n} > 0$.

Khi đó $n > 1$ và bằng cách áp dụng hệ quả của Kết quả 1 ta thu được

$$(\text{per}(A))^2 = \text{per}(a_1, \dots, a_{n-1}, a_n)^2 \geq \text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \text{per}(a_1, \dots, a_{n-2}, a_n, a_n).$$

Mặt khác, $\text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) = \sum_j a_{(n-1),j} \text{per}(A_{n,j}) > \sum_j a_{(n-1),j} \text{per}(A) = \text{per}(A)$,

$$\text{và } \text{per}(a_1, \dots, a_{n-2}, a_n, a_n) = \sum_j a_{n,j} \text{per}(A_{(n-1),j}) \geq \sum_j a_{n,j} \text{per}(A) = \text{per}(A),$$

nên $(\text{per}(A))^2 > \text{per}(A)^2$, mâu thuẫn.

Do đó $\text{per}(A_{nn}) = \text{per}(A)$.

Bổ đề 14. Nếu A là ma trận minimal với $a_{ij} > 0$ hầu hết tại mọi nơi ngoại trừ khi $i = n$ thì $a_{ij} = \frac{1}{n}$ với mọi i, j .

Chứng minh. Từ Bổ đề 13 ta có $\text{per}(a_1, \dots, a_{n-2}, a_n, a_n) = \sum_j a_{n,j} \text{per}(A) = \text{per}(A)$, tương tự

$\text{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) = \text{per}(A)$. Áp dụng điều kiện xảy ra dấu “=” nêu ở Kết quả 1 ta suy ra $a_n = \lambda a_{n-1}$ với một số λ nào đó. Do A là ma trận song ngẫu nhiên nên ta phải có $\lambda = 1$ và do đó $a_n = a_{n-1}$. Lập luận tương tự ta suy ra được rằng tất cả các hàng đều bằng nhau và tương tự tất cả các cột đều bằng nhau. Điều này dẫn đến rằng tất cả các phần tử của ma trận A đều bằng $\frac{1}{n}$.

Kết quả 3. Nếu A là ma trận minimal thì $a_{ij} = \frac{1}{n}$ với mọi i, j , và khi đó $\text{per}(A) = \frac{n!}{n^n}$

Chứng minh. Đặt B là ma trận thu được từ A bằng cách thay một hàng a_i bởi một hàng a_k khác; khi đó $\text{per}(B) = \sum_j a_{k,j} \text{per}(A_{ij}) = \text{per}(A)$. Tương tự nếu ta gọi C là ma trận thu được từ A bằng cách thay a_k bởi a_i thì ta cũng nhận được $\text{per}(C) = \text{per}(A)$. Tuy rằng chúng ta không thể biết được chính xác B và C có phải là các ma trận song ngẫu nhiên hay không nhưng ta có thể khẳng định chắc chắn được rằng ma trận

$D = \frac{1}{2}(B + C)$ sẽ là một ma trận song ngẫu nhiên và $\text{per}(D) = \frac{1}{4}(\text{per}(B) + 2\text{per}(A) + \text{per}(C)) = \text{per}(A)$

nên D cũng là một ma trận minimal.

Sau một số hữu hạn bước thực hiện các quá trình “trung bình hóa” như ta vừa trình bày khi xây dựng ma trận D từ ma trận A ta thu được ma trận minimal E có cùng hàng cuối cùng như ma trận A . Từ cách xây dựng vừa trình bày ta thấy rằng $e_{ij} = 0$ chỉ khi $i = n$ hoặc $a_{1j} = \dots a_{(n-1),j} = 0$. Do ma trận A không thể là decomposable nên ta không thể có $a_{1j} = \dots a_{(n-1),j} = 0$. Do vậy E là một ma trận minimal thỏa mãn tất cả các

điều kiện nêu trong Bổ đề 14 nên ta dễ dàng suy ra hàng cuối cùng a_n của nó sẽ bằng $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$. Lập luận

tương tự ta cũng nhận thấy tất cả các hàng khác của A cũng có cùng giá trị này. Điều này kết thúc quá trình chứng minh.

Tài liệu tham khảo

[1]. Marshall Hall, Combinatorial Theory, First Edition

[2]. Trương Phước Nhân, Permanent, 03/03/2018