Permanent của ma trận

Trương Phước Nhân, 13/07/2018

1. Định nghĩa và các tính chất cơ bản

Cho trước ma trận $A=\left(a_{i,j}\right)_{i=1,\dots,n;\ j=1,\dots,m}$ $\left(n\le m\right)$ với các phần tử nằm trên một vành $\mathbb{F}$. Permanent của ma trận A là một hàm nhận giá trị trên vành $\mathbb{F}$ được xác định bởi hệ thức $\operatorname{per}A=\sum_{j_1,\dots,j_n} a_{1j_1}\dots a_{nj_n}$, trong đó tổng được lấy trên tất cả các chỉnh hợp chập n của m chỉ số phân biệt cho trước. Nếu $m=n$ thì tổng trên được lấy trên tất cả các hoán vị của các phần tử $1,2,\dots,n$ và khi đó permanent của ma trận có dạng tương tự như định thức của một ma trận vuông về mặt hình thức.

Sau đây là một số tính chất cơ bản của permanent có dạng tương tự về mặt hình thức với định thức:

(1) Nếu ma trận A cấp $n\times m$ với $n\le m$ có chứa một hàng gồm toàn số 0 thì $\operatorname{per}A$.

(2) Thực hiện phép toán nhân một hàng nào đó trong số n hàng của một ma trận A cấp $n\times m\left(n\le m\right)$ với một phần tử $c\in\mathbb{F}$ biến $\operatorname{per}A$ thành $c\operatorname{per}A$.

(3) Việc đổi chỗ các hàng và các cột của một ma trận A không làm thay đổi permanent của nó; nói cách khác, nếu π_1 và π_2 lần lượt là các ma trận hoán vị bậc n và m thì $\operatorname{per}\left(\pi_1 A\pi_2\right)=\operatorname{per}A$.

(4) Ký hiệu A_{ij} là ma trận thu được từ ma trận A bằng cách xóa đi hàng thứ i và cột thứ j, còn A'_{ij} là ma trận thu được từ ma trận A bằng cách thay thế phần tử a_{ij} bởi 0. Khi đó $\operatorname{per}A=\operatorname{per}A'_{ij}+a_{ij}\operatorname{per}A_{ij}$.

Thật vậy, hệ thức trên thu được bằng cách phân các số hạng trong công thức tính $\operatorname{per}A$ thành hai lớp sao cho một lớp gồm tất cả các số hạng có chứa phần tử a_{ij} và lớp kia chứa tất cả các số hạng còn lại. Tổng tất cả các số hạng chứa a_{ij} bằng $a_{ij}\operatorname{per}A_{ij}$ còn tổng còn lại thì bằng $\operatorname{per}A'_{ij}$.

Hệ thức vừa nêu thường được gọi là khai triển của permanent theo phần tử a_{ij}.

(5) Nếu $a_{ij_1},\dots,a_{ij_k}$ là các phần tử khác không của hàng thứ i thì $\operatorname{per}A=\sum_{v=1}^{k} a_{ij_v}\operatorname{per}A_{ij_v}$.

Hệ thức trên được gọi là phân tích của permanent theo hàng thứ i. Nó thu được bằng cách áp dụng liên tiếp khai triển của permanent theo các phần tử.

(6) Đặt $A\left[i_1,\dots,i_r\mid j_1,\dots,j_s\right]$ là ma trận con của A được xác định bởi giao giữa các hàng được đánh số $i_1,\dots,i_r$ và các cột được đánh số $j_1,\dots,j_s$; $A\left(i_1,\dots,i_r\mid j_1,\dots,j_s\right)$ là ma trận thu được từ A bằng cách xóa đi các hàng $i_1,\dots,i_r$ và các cột $j_1,\dots,j_s$.

Khi đó, với $r\le s, n-r\le m-s$, ta có

$$\operatorname{per}A=\sum_{j_1,\dots,j_s}\operatorname{per}A\left[i_1,\dots,i_r\mid j_1,\dots,j_s\right]\operatorname{per}A\left(i_1,\dots,i_r\mid j_1,\dots,j_s\right),$$ trong đó tổng được lấy theo tất cả các tổ hợp chập s của m phần tử $1,\dots,m$.

Hệ thức này thường được gọi là công thức khai triển Laplace của permanent.

Cho A là một ma trận vuông khối với các khối vuông $A_1,\dots,A_k$ nằm dọc trên đường chéo chính còn tất cả các phần tử khác của ma trận A đều bằng không, ký hiệu $A=\operatorname{diag}\left(A_1,\dots,A_k\right)$.

Từ công thức khai triển Laplace cho permanent ta suy ra $\operatorname{per}A=\operatorname{per}A_1\dots\operatorname{per}A_k$.

Như vừa trình bày ở phần trên, permanent có một số tính chất tương tự với định thức nhưng ta cần lưu ý là trong trường hợp tổng quát hệ thức $\det AB=\det A\det B$ không đúng cho permanent. Điều này có nghĩa rằng khác với khái niệm định thức, permanent của một ma trận với các hàng hoặc các cột phụ thuộc tuyến tính có thể không cần bằng không.

2. Phương pháp tính permanent

Trong mục này ta nêu ra một số ví dụ cho phương pháp tính permanent của các ma trận vuông sở hữu một số tính chất đối xứng đặc biệt. Tính đối xứng này cho phép ta chuyển bài toán tính permanent về dạng đơn giản hơn hoặc thiết lập hệ thức truy hồi để xác định giá trị của permanent.

2.1. Permanent của tổ hợp tuyến tính của hai ma trận hoán vị

Cho φ_1 và φ_2 là các hoán vị bậc n và π_1 và π_2 là các ma trận hoán vị tương ứng với φ_1 và φ_2.

Khi đó ma trận $\alpha\pi_1+\beta\pi_2$, trong đó α,β là các số thực, là một tổ hợp tuyến tính của các ma trận π_1 và π_2.

Ta sẽ chứng minh rằng

$$\operatorname{per}(\alpha\pi_1+\beta\pi_2)=\prod_{i=1}^{k}\left(\alpha^{l_i}+\beta^{l_i}\right), \text{ trong đó } l_1,\dots,l_k \text{ là độ dài các chu trình của hoán vị } \varphi_1^{-1}\varphi_2.$$

Giả sử hoán vị $\varphi_1^{-1}\varphi_2$ có dạng $\varphi_1^{-1}\varphi_2=(a_1,\dots,a_{l_1})(b_1,\dots,b_{l_2})\dots(c_1,\dots,c_{l_k})$ và xét hoán vị

$$\psi=\begin{pmatrix} a_1 & \dots & a_{l_1} & b_1 & \dots & b_{l_2} & \dots & c_1 & \dots & c_{l_k} \\ 1 & \dots & l_1 & l_1+1 & \dots & l_1+l_2 & \dots & n-l_k+1 & \dots & n \end{pmatrix}$$

Dễ dàng kiểm tra được rằng $\psi^{-1}\varphi_1^{-1}\varphi_2\psi=(1,\dots,l_1)(l_1+1,\dots,l_1+l_2)\dots(n-l_k+1,\dots,n)$.

Ký hiệu π là ma trận hoán vị tương ứng với hoán vị ψ và đặt $\overline{\pi}=\pi^{-1}\pi_1^{-1}\pi_2\pi$.

Khi đó: $\operatorname{per}(\alpha\pi_1+\beta\pi_2)=\operatorname{per}(\alpha E+\beta\overline{\pi})$, trong đó $\overline{\pi}=\operatorname{diag}(\overline{\pi_1},\dots,\overline{\pi_k})$ với $\overline{\pi_i}$ là ma trận hoán vị cấp l_i tương ứng với chu trình $(1,\dots,l_i)$, $i=1,\dots,k$ còn E là ma trận đơn vị.

Do đó: $\operatorname{per}(\alpha\pi_1+\beta\pi_2)=\operatorname{perdiag}(A_1,\dots,A_k)$, trong đó A_i, $i=1,\dots,k$, là ma trận vuông cấp l_i có dạng

$$A_i=\begin{pmatrix} \alpha & \beta & & & & 0 \\ & \alpha & \beta & & & \\ & & \alpha & \beta & & \\ & & & \dots & & \\ 0 & & & & \alpha & \beta \\ \beta & & & & & \alpha \end{pmatrix}.$$

Khai triển $\operatorname{per}A_i$ theo phần tử nằm ở góc dưới bên tay trái ta nhận được $\operatorname{per}A_i=\alpha^{l_i}+\beta^{l_i}$, $i=1,\dots,k$.

Tóm lại, $\operatorname{per}(\alpha\pi_1+\beta\pi_2)=\prod_{i=1}^{k}\left(\alpha^{l_i}+\beta^{l_i}\right)$.

2.2. Permanent của tổ hợp tuyến tính của ba ma trận hoán vị

Cho C là ma trận hoán vị tương ứng với chu trình c bậc n với $c=(1,\dots,n)$.

Xét tổ hợp tuyến tính $Q_n=\alpha E+\beta C+\gamma C^2$, trong đó α,β,γ là các số thực và E là một ma trận đơn vị.

Ta sẽ chứng minh rằng, với $n>3$, $\operatorname{per}(\alpha E+\beta C+\gamma C^2)=r_1^n+r_2^n+\alpha^n+\gamma^n$, trong đó r_1,r_2 là các phân biệt nghiệm của phương trình $x^2-\beta x-\alpha\gamma=0$.

$$\text{Đặt } D_n=\operatorname{per}\begin{pmatrix} \beta & \gamma & & & & 0 \\ \alpha & \beta & \gamma & & & \\ & \alpha & \beta & \gamma & & \\ & & & \dots & & \\ & & & \alpha & \beta & \gamma \\ 0 & & & & \alpha & \beta \end{pmatrix}.$$

Khai triển D_n theo hàng đầu tiên, ta nhận được hệ thức truy hồi sau $D_n=\beta D_{n-1}+\alpha\gamma D_{n-2}$.

Đây là một phương trình sai phân tuyến tính hệ số hằng với phương trình đặc trưng $x^2-\beta x-\alpha\gamma=0$.

Do đó, $D_n=C_1r_1^n+C_2r_2^n$, trong đó r_1,r_2 là các nghiệm của phương trình $x^2-\beta x-\alpha\gamma=0$.

Các hệ số C_1 và C_2 được xác định từ điều kiện ban đầu $D_0=1, D_1=\beta$ và do đó ta tính được

$$C_1=\frac{r_1}{\mu},\ C_2=\frac{r_2}{\mu}, \text{ với } \mu=\sqrt{\beta^2+4\alpha\gamma}\neq 0.$$

Do đó, $D_n = \frac{1}{\mu}\left(r_1^{n+1} - r_2^{n+1}\right)$, trong đó $r_1 = \frac{\beta+\mu}{2}$ và $r_1 = \frac{\beta-\mu}{2}$.

Khai triển $\text{per}\left(\alpha E + \beta C + \gamma C^2\right)$ theo ba phần tử khác không bắt đầu từ phần tử nằm ở góc dưới cùng bên tay trái, ta nhận được hệ thức truy hồi sau

$$\text{per}\left(\alpha E + \beta C + \gamma C^2\right) = \beta D_{n-1} + 2\alpha\gamma D_{n-2} + \alpha^n + \gamma^n.$$

Thay D_n bởi biểu thức tính được vào khai triển của $\text{per}\left(\alpha E + \beta C + \gamma C^2\right)$ ta nhận được khẳng định ban đầu.

Lưu ý: Trong trường hợp $\mu = 0$, ta tính được $D_n = (n+1)(\beta/2)^n$ và bằng phương pháp tương tự ta cũng có thể xác định được biểu thức tính của $\text{per}\left(\alpha E + \beta C + \gamma C^2\right)$.

3. Đánh giá permanent

3.1. Đánh giá cho permanent của các ma trận không âm

Từ công thức Ryser và bất đẳng thức Bonferroni suy ra rằng với mọi ma trận không âm $A = \left(a_{ij}\right)_{n\times n}$, ta luôn có

$$\text{per}A \le \prod_{i=1}^{n}\sum_{j=1}^{n} a_{ij}.$$

Đánh giá vừa nêu là rất thô và trong một số trường hợp nhất định ta có thể cải thiện đánh giá trên.

Nhắc lại rằng một ma trận không âm $A = \left(a_{ij}\right)_{n\times n}$ được gọi là irreducible nếu như không tồn tại một ma trận hoán vị π sao cho $\pi^{-1}A\pi = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}$, trong đó A_1, A_3 là các ma trận vuông nào đó.

Từ Định lý Perron – Frobenius, ma trận không âm A có một giá trị riêng dương r là một nghiệm đơn nào đó của phương trình đặc trưng và vector riêng $x = \left(x_1, ..., x_n\right)$ tương ứng với r có tất cả các thành phần là các số thực dương.

Điều này có nghĩa rằng $Ax^T = rx^T$.

Khai triển hệ thức trên ta suy ra rằng $\sum_{j=1}^{n} a_{ij}x_j = rx_i, i = 1, ..., n$.

Bằng cách lấy tích các phương trình trên, ta nhận được $\prod_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j = r^n\prod_{i=1}^{n} x_i$.

Từ định nghĩa của permanent $\text{per}A = \sum_{i_1,...,i_n} a_{1i_1}...a_{ni_n}$, ta suy ra được đánh giá sau

$$\prod_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j \ge \text{per}A\prod_{i=1}^{n} x_i, \text{ với mọi số thực dương } x_1, ..., x_n.$$

Kết hợp hai kết quả vừa thu nhận được ở trên ta suy ra rằng $\text{per}A \le r^n$.

Nếu ma trận irreducible A là song ngẫu nhiên thì $0 < \text{per}A \le 1$, vì một ma trận song ngẫu nhiên thì luôn có giá trị riêng dương bằng 1, và chặn trên của đánh giá đặt được cho các ma trận hoán vị.

3.2. Đánh giá cho permanent của các $(0,1)-$ma trận

Trong phần này của bài viết ta nêu ra các đánh giá cho chặn trên và chặn dưới của các $(0,1)-$ma trận dựa vào tổng các phần tử trên các hàng của nó.

Đầu tiên ta đưa ra một chặn dưới được đề xuất bởi Minc.

Xét một $(0,1)-$ma trận A sao cho $\sum_{j=1}^{n} a_{ij} = r_i \ge 0, i = 1, ..., n$.

Khi đó, $\text{per}A \ge \prod_{i=1}^{n} \max\left\{r_i - n + i, 0\right\}$.

Thật vậy, ta chứng minh đánh giá trên bằng phương pháp quy nạp theo n.

Với $r_n = 0$, đánh giá là hiển nhiên. Nếu $r_n > 0$, không mất tính tổng quát ta có thể giả sử rằng $a_{nj} = 1$ với $j = 1, ..., r_n$.

Khai triển $\operatorname{per}A$ theo hàng cuối cùng và sử dụng giả thiết quy nạp theo n ta nhận được

$$\begin{aligned}\operatorname{per}A=\sum_{j=1}^{r_n}\operatorname{per}A(n\,|\,j)&\geq\sum_{j=1}^{r_n}\prod_{i=1}^{n-1}\max\{r_i-a_{ij}+i-n+1,0\}\\&\geq\sum_{j=1}^{r_n}\prod_{i=1}^{n-1}\max\{r_i+i-n,0\}\\&=\prod_{i=1}^{n}\max\{r_i-n+i,0\}.\end{aligned}$$

Tiếp theo ta sẽ chỉ ra một chặn dưới cho các $(0,1)-$ma trận cũng được đề xuất bởi tác giả Minc nhưng chỉ được chứng minh chặt chẽ bởi nhà toán học Bregman:

$$\operatorname{per}A\leq\prod_{i=1}^{n}(r_i!)^{\frac{1}{r_i}}.$$

Để thực hiện được điều vừa trình bày ở trên ta cần đến hai kết quả phụ sau:

Bổ đề 1. Với mọi số thực không âm $v_1,...,v_q$ ta luôn có $(v_1+...+v_q)^{v_1+...+v_q}\leq v_1^{v_1}...v_q^{v_q}q^{v_1+...+v_q}$, trong đó ta đã quy ước rằng $0^0=1$.

Chứng minh. Thật vậy, bất đẳng thức $(v_1+...+v_q)^{v_1+...+v_q}\leq v_1^{v_1}...v_q^{v_q}q^{v_1+...+v_q}$ cần chứng minh tương đương với đánh giá $\dfrac{v_1+...+v_q}{q}\ln\left(\dfrac{v_1+...+v_q}{q}\right)\leq\dfrac{v_1\ln v_1+...+v_q\ln v_q}{q}$. Tính đúng đắn của kết quả này là một hệ quả đơn giản từ tính lồi của hàm $y=x\ln x$.

Bổ đề 2. Cho $A=(a_{ij})_{n\times n}$ là một $(0,1)-$ma trận và S là tập tất cả các hoán vị s bậc n sao cho $\prod_{i=1}^{n}a_{i,s(i)}=1$.

Khi đó, $\prod_{s\in S}r_i=r_i^{\operatorname{per}A},\ i=1,...,n,$

$$\prod_{s\in S}\operatorname{per}A(i\,|\,s(i))=\prod_{j:\,a_{ij}=1}(\operatorname{per}A(i\,|\,j))^{\operatorname{per}A(i|j)},i=1,...,n.$$

Chứng minh. Hệ thức thứ nhất suy trực tiếp từ định nghĩa của permanent.

Với hai chỉ số i,j bất kỳ, ta có $|\{s:s(i)=j,a_{ij}=1,s\in S\}|=\operatorname{per}A(i\,|\,j)$, nên do đó

$$\prod_{s\in S}\operatorname{per}A(i\,|\,s(i))=\prod_{j:\,a_{ij}=1}\ \prod_{s:\,s(i)=j}\operatorname{per}A(i\,|\,j)=\prod_{j:\,a_{ij}=1}(\operatorname{per}A(i\,|\,j))^{\operatorname{per}A(i|j)}.$$

Trở lại với vấn đề ban đầu, ta sẽ chứng minh khẳng định bằng quy nạp theo n.

Với $n=1$, đánh giá là hiển nhiên.

Đặt $v_k=a_{ik}\operatorname{per}A(i\,|\,k)$ với mọi chỉ số k sao cho $a_{ik}=1$.

Áp dụng Bổ đề 1 cho $q=r_i$ vào vế phải của hệ thức sau

$$(\operatorname{per}A)^{\operatorname{per}A}=\left(\sum_{j:\,a_{ij}=1}\operatorname{per}A(i\,|\,j)\right)^{\sum_{j:\,a_{ij}=1}\operatorname{per}A(i|j)},$$

ta nhận được đánh giá $$(\operatorname{per}A)^{\operatorname{per}A}\leq r_i^{\operatorname{per}A}\prod_{j:\,a_{ij}=1}(\operatorname{per}A(i\,|\,j))^{\operatorname{per}A(i|j)}.$$

Áp dụng Bổ đề 2 ta nhận được $(\operatorname{per}A)^{\operatorname{per}A}\leq\prod_{s\in S}r_i\operatorname{per}A(i\,|\,s(i))$.

Để thực hiện đánh giá $\operatorname{per}A(i\,|\,s(i))$ ta sẽ cần đến giả thiết quy nạp. Ký hiệu r_k' là tổng các phần tử của hàng thứ k của ma trận $A(i\,|\,s(i))$, khi đó $r_k'=\begin{cases}r_k-1\ ,a_{k,s(i)}=1,\\ r_k\quad\ ,a_{k,s(i)}=0.\end{cases}$

Áp dụng giả thiết quy nạp ta nhận được $\operatorname{per}A(i\,|\,s(i)) \le \prod_{k\ne i:\, a_{k,s(i)}=1} \left((r_k-1)!\right)^{\frac{1}{r_k-1}} \prod_{k\ne i:\, a_{k,s(i)}=0} (r_k!)^{\frac{1}{r_k}}$.

Bằng cách lấy tích theo vế với vế đánh giá trên khi $i=1,\ldots,n$ và đổi thứ tự tính tích giữa k và i, ta được

$$\prod_{i=1}^{n} \operatorname{per}A(i\,|\,s(i)) \le \prod_{k=1}^{n} \prod_{i\ne k:\, a_{k,s(i)}=1} \left((r_k-1)!\right)^{\frac{1}{r_k-1}} \prod_{i\ne k:\, a_{k,s(i)}=0} (r_k!)^{\frac{1}{r_k}} .$$

Với k bất kỳ và $s\in S$, $\left|\left\{i : a_{k,s(i)}=0\right\}\right| = n-r_k$,

$$\left|\left\{i : a_{k,s(i)}=1\right\}\right| = r_k-1, i\ne k .$$

Từ phân tích vừa nêu ta thu nhận được

$$\prod_{i=1}^{n} \operatorname{per}A(i\,|\,s(i)) \le \prod_{k=1}^{n} (r_k!)^{\frac{n-r_k}{r_k}} (r_k-1)!.$$

Mặt khác, bằng cách lấy tích theo vế với vế của hệ thức $(\operatorname{per}A)^{\operatorname{per}A} \le \prod_{s\in S} r_i \operatorname{per}A(i\,|\,s(i))$ khi cho $i=1,\ldots,n$ và sử dụng đánh giá $\prod_{i=1}^{n} \operatorname{per}A(i\,|\,s(i)) \le \prod_{k=1}^{n} (r_k!)^{\frac{n-r_k}{r_k}} (r_k-1)!$ ta thu được

$$(\operatorname{per}A)^{n\operatorname{per}A} \le \prod_{s\in S} \prod_{k=1}^{n} (r_k!)^{\frac{n}{r_k}} .$$

Do $|S| = \operatorname{per}A$ nên từ đánh giá trên ta cũng nhận ra rằng $(\operatorname{per}A)^{n\operatorname{per}A} \le \left(\prod_{k=1}^{n} (r_k!)^{\frac{n}{r_k}}\right)^{\operatorname{per}A}$

$$\Rightarrow \operatorname{per}A \le \prod_{i=1}^{n} (r_i!)^{\frac{1}{r_i}} .$$

Nhận xét:

Bằng cách áp dụng bất đẳng thức AM-GM ta dễ dàng chứng minh được $(r_i!)^{\frac{1}{r_i}} \le \frac{r_i+1}{2}$.

Thật vậy, $(r_i!)^{\frac{1}{r_i}} \le \frac{1}{r_i}\sum_{k=1}^{r_i} k = \frac{1}{r_i}\frac{r_i(r_i+1)}{2}$.

Do đó, $\operatorname{per}A \le \prod_{i=1}^{n} \frac{r_i+1}{2}$.

Tài liệu tham khảo:

[1]. Trương Phước Nhân, Định lý Van der Waerden, 08/05/2018.
[2]. Trương Phước Nhân, Permanent, 03/03/2018.
[3]. Vladimir N. Sachkov, Combinatorial Methods in Discrete Mathematics, Cambridge University Press.