European Journal of Combinatorics 31 (2010) 1887–1891

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

The permanent of a square matrix

David G. Glynn

CSEM, Flinders University, P.O. Box 2100, Adelaide, South Australia 5001, Australia

ARTICLE INFO

Article history:
Received 10 December 2009
Accepted 25 January 2010
Available online 10 February 2010

ABSTRACT

We investigate the permanent of a square matrix over a field and calculate it using ways different from Ryser's formula or the standard definition. One formula is related to symmetric tensors and has the same efficiency $O(2^m m)$ as Ryser's method. Another algebraic method in the prime characteristic case uses partial differentiation.

1. Introduction

Let M be the set $\{1, \ldots, m\}$, $(m \in \mathbb{Z}^+)$. The *symmetric group* S_m is the group of all $m!$ permutations of M. The *sign* of a permutation $\gamma \in S_m$ is

$$\operatorname{sgn}(\gamma) := \begin{cases} 1 & \text{for } \gamma \text{ even} \\ -1 & \text{for } \gamma \text{ odd.} \end{cases}$$

Let $A = (a_{ij})$ be an $m \times m$ matrix over a field F. The *determinant* of A is a polynomial of degree m: $\det(A) := \sum_{\gamma \in S_m} \operatorname{sgn}(\gamma) \prod_{i=1}^m a_{i,\gamma(i)}$.

The *permanent* of A is a similar polynomial of degree m: $\operatorname{per}(A) := \sum_{\gamma \in S_m} \prod_{i=1}^m a_{i,\gamma(i)}$. Note that if F has characteristic two, the permanent equals the determinant. If the x_i's are commuting indeterminants $\operatorname{per}(A)$ is the coefficient of $x_1 \ldots x_n$ in the polynomial $\prod_{j=1}^m \sum_{i=1}^m x_i \cdot a_{i,j}$.

H.J. Ryser found an alternative method to evaluate $\operatorname{per}(A)$ by the method of inclusion and exclusion. See [1,5] for some of the theories of permanents. See [2] for some elementary number theories that we use here.

For $S \subseteq M$ let $\bar{S}$ be the complementary subset $M \setminus S$. The cardinality of the subset S is denoted $|S|$. Then

$$\operatorname{per}(A) = \sum_{S \subseteq M} (-1)^{|S|} \prod_{j=1}^m \sum_{i \in \bar{S}} a_{i,j}.$$

E-mail addresses: davidg@csem.flinders.edu.au, dglynn@me.com.

0195-6698/$ – see front matter
doi:10.1016/j.ejc.2010.01.010

With minor modifications (such as using a Gray code to minimise the number of additions) Ryser's formula is often used to calculate the permanent on a computer. By counting multiplications it has efficiency $O(2^m m)$.

Here are three ways to calculate per(A) for a general 3×3 matrix

$$A := \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}.$$

The last two methods are faster for larger general matrices.

Example 1.1. The classical formula using all the permutations in S_3 is

$$\text{per}(A) = aei + bfg + cdh + afh + bdi + ceg.$$

Ryser's method gives

$$\begin{aligned}\text{per}(A) &= (a+b+c)(d+e+f)(g+h+i) - (a+b)(d+e)(g+h) \\ &\quad - (a+c)(d+f)(g+i) - (b+c)(e+f)(h+i) + adg + beh + cfi.\end{aligned}$$

Theorem 2.1 gives

$$\begin{aligned}2^2\text{per}(A) &= (a+b+c)(d+e+f)(g+h+i) - (a-b+c)(d-e+f)(g-h+i) \\ &\quad - (a+b-c)(d+e-f)(g+h-i) + (a-b-c)(d-e-f)(g-h-i).\end{aligned}$$

2. The permanent and the polarization identity

There is another formula that has a similar computational efficiency $O(2^m m)$ to Ryser's formula. It is related to the polarization identity for symmetric tensors.

Theorem 2.1. *Let $A = (a_{ij})$ be an $m \times m$ matrix over a field F of characteristic not two. Then*

$$\text{per}(A) = \left[\sum_{\delta} \left(\prod_{k=1}^{m} \delta_k\right) \prod_{j=1}^{m} \sum_{i=1}^{m} \delta_i a_{ij}\right] \Big/ 2^{m-1},$$

where the outer sum is over all 2^{m-1} vectors $\delta = (\delta_1, \ldots, \delta_m) \in \{\pm 1\}^m$, with $\delta_1 = 1$.

Proof. Consider a general monomial g in the right-hand side (times 2^{m-1}). It has degree m in the a_{ij}'s, one in each column j of A. Let the number of a_{ij}'s of g occurring in row i be λ_i. Clearly $\sum_{i=1}^{m} \lambda_i = m$ and $0 \le \lambda_i \le m$. It is seen that the coefficient c of g will be $\sum_\delta \prod_{i=1}^{m} (\delta_i)^{\lambda_i+1} = \prod_{i=2}^{m} \sum_{\delta_i \in \{\pm 1\}} (\delta_i)^{\lambda_i+1}$. Now

$$\sum_{\delta_i \in \{\pm 1\}} (\delta_i)^{\lambda_i+1} = \begin{cases} 0 & \lambda_i \text{ even} \\ 2 & \lambda_i \text{ odd.} \end{cases}$$

Thus c is zero unless λ_i is odd for all $i > 1$, in which case

$$\sum_{i=2}^{m} \lambda_i \begin{cases} = m-1 & \lambda_i = 1, \forall 2 \le i \le m \\ \ge m+1 & \text{otherwise.} \end{cases}$$

But the latter is impossible and so only non-zero coefficients for monomials on the right-hand side appear when all the λ_i's are equal to one, and in this case $c = 2^{m-1}$. Such monomials correspond to products $g = \prod_{i=1}^{m} a_{i,\sigma(i)}$, where $\sigma \in S_m$. $\square$

Let $S^m(V) := \text{Sym}(V^{\otimes m})$ be the space of symmetric m-tensors over a vector space V. See [4, Section B.2.3]. It is the image of the projection

$$\text{Sym}(v_1 \otimes \cdots \otimes v_m) := \sum_{s \in S_m} \sigma_m(s)(v_1 \otimes \cdots \otimes v_m)/m!,$$

where $\sigma_m(s)$ permutes the positions of the factors in the tensor product.

The *polarization identity* shows that $S^m(V)$ is spanned by the m-fold tensor products of type $v^{\otimes m} = \otimes_{j=1}^m v$, for $v \in V$, and it has the formula:

$$\mathrm{Sym}(v_1 \otimes \cdots \otimes v_m) = \sum_{\delta} \left(\prod_{k=1}^m \delta_k \right) \bigotimes_{j=1}^m \left(\sum_{i=1}^m \delta_i v_i \right) / (m! 2^{m-1}),$$

where the outer sum is over all 2^{m-1} vectors $\delta = (\delta_1, \ldots, \delta_m) \in \{\pm 1\}^m$, with $\delta_1 = 1$.

A correspondence between this formula and Theorem 2.1 is given by the transformation

$$v_{u_1} \otimes \cdots \otimes v_{u_m} \leftrightarrow a_{u_1 1} \ldots a_{u_m m}.$$

This is related to the well-known isomorphism between the polynomial algebra over a field, and the symmetric algebra over a vector space. The latter (if the characteristic over the field does not interfere) can be identified with the space of symmetric tensors.

3. The permanent from partial derivatives

Theorem 3.1. *Let $A = (a_{ij})$ be an $m \times m$ matrix over a field of characteristic p. Then there holds*

$$\mathrm{per}(A) = (-1)^m \left(\frac{\partial^m}{\partial a_{11} \ldots \partial a_{mm}} \right)^{p-2} \det(A)^{p-1}.$$

Proof. The polynomial $f := \det(A)^{p-1}$ is homogeneous of degree $m(p-1)$ in the m^2 variables a_{ij}. To obtain a non-zero monomial in these variables from the operator

$$\left(\frac{\partial^m}{\partial a_{11} \cdots \partial a_{mm}} \right)^{p-2} = \prod_{i=1}^m \frac{\partial^{p-2}}{\partial a_{ii}^{p-2}}$$

that acts upon f, we must consider monomials in f that are of the form $K := \prod_{i=1}^m a_{ii}^{p-2} \cdot k$, where k is a homogeneous polynomial of degree m. Since $\det(A)$ is the sum of products of the form $\prod_{i=1}^m a_{i,\gamma(i)}$, it is clear that k must be of a similar form, corresponding to a unique permutation γ of S_m. Let γ have r even cycles of sizes $a_1, \ldots, a_r$, $(a_i \geq 2)$, s odd cycles of sizes $b_1, \ldots, b_s$, $(b_i \geq 3)$, and c cycles of size one, i.e. fixed points. Then we note that $m \equiv s + c \pmod 2$, since the even cycles can be neglected when we calculate the length of the permutation mod 2. Also, γ is an even permutation if and only if r is even.

When k is produced by the above partial differentiation from K there is a factor of $((p-2)!)^{r+s} \cdot ((p-1)!)^c \equiv (-1)^c$ in $\mathrm{GF}(p)$. This is because a cycle of size one in γ corresponds to an a_{ii}^{p-1} in K, and differentiates $p-2$ times to $(p-1)! a_{ii}$. Otherwise, $a_{ii}^{p-2} a_{jk}$ differentiates to $(p-2)! a_{jk}$, and $(p-2)! \equiv 1$ while $(p-1)! \equiv -1 \pmod p$. Next, each function from $\{1, \ldots, r+s\} \to \{1, \ldots, p-1\}$ corresponds to a way that K can appear as a polynomial in the product of $\det(A)$, $p-1$ times. This is because the non-diagonal a_{ij}'s of k in each non-trivial cycle must be assigned to one of the $p-1$ permutations of the A's. The number of these functions is $(p-1)^{r+s} \equiv (-1)^{r+s}$ in $\mathrm{GF}(p)$. In addition, each of these ways of producing k gives the same sign as γ, since the $p-1$ permutations from the determinants have a product that is γ, and so the signs multiply to give $\mathrm{sgn}(\gamma) = (-1)^r$.

Hence the total coefficient of k in the partial derivative of K and hence of $\det(A)^{p-1}$ is $(-1)^r \cdot (-1)^c \cdot (-1)^{r+s} \equiv (-1)^{s+c} \equiv (-1)^m \pmod p$. Thus the formula in the theorem is correct. □

There is a more general result, with a different proof. These algebraic methods of calculating the permanent are symbolic, and so cannot be compared directly in efficiency with the methods of Ryser and Theorem 2.1.

Theorem 3.2. *Let $A = (a_{ij})$ be an $m \times m$ matrix over a field of characteristic p. Let $B = (b_{ij})$ be any $m \times m$ integer matrix with $0 \leq b_{ij} \leq p-2$ and with every row and column sum $p-2$. Then there holds*

$$\mathrm{per}(A) = (-1)^m \prod_{i=1}^m \prod_{j=1}^m \left(\frac{\partial^{b_{ij}}}{\partial a_{ij}^{b_{ij}}} \right) \det(A)^{p-1}.$$

Proof. Use the theory of the p-modular hyperdeterminant $\det_p$. By [3, Theorem 4.1] $\det(A)^{p-1} = \det_p(A)$, and by the formula for $\det_p$ [3, Theorem 6.1], this implies that $(-1)^m \det(A)^{p-1}$ is the sum of all terms $\prod_{i=1}^m \prod_{j=1}^m a_{ij}^{e_{ij}}/e_{ij}!$ in the m^2 indeterminates a_{ij}, where $E = (e_{ij})$ is a general non-negative integer matrix with row and column sums $p-1$. The result of differentiating this term with respect to the fixed matrix B, as above, with row and column sums $p-2$, will be zero unless $b_{ij} \le e_{ij}$ for all i, j. In that case, from the row and column sums, this can only happen if $E - B$ is a permutation matrix and so $e_{ij} = b_{ij}$ or $b_{ij} + 1$, for all i, j. Then

$$\left(\frac{\partial^{b_{ij}}}{\partial a_{ij}^{b_{ij}}}\right) a_{ij}^{e_{ij}}/e_{ij}! = 1 \quad \text{or } a_{ij} \text{ respectively.}$$

The a_{ij}'s occur in the general permutation $E - B$, and hence we obtain $(-1)^m$ times $\mathrm{per}(A)$. □

Corollary 3.3. *Let $A = (a_{ij})$ be an $m \times m$ matrix over a field of characteristic three. Then*

$$\mathrm{per}(A) = (-1)^m \frac{\partial^m}{\partial a_{11} \dots \partial a_{mm}} \det(A)^2.$$

Example 3.4.

$$\begin{aligned}
\mathrm{per}\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= (-1)^2 \frac{\partial^2}{\partial a \partial d} \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 \\
&= \frac{\partial^2}{\partial a \partial d}(ad - bc)^2 = \frac{\partial}{\partial a} 2a(ad - bc) \\
&= 2(ad - bc) + 2ad = ad + bc \pmod 3.
\end{aligned}$$

Theorem 3.5. *Let $A = (a_{ij})$ be an $m \times m$ matrix over a field of characteristic three. Then*

$$\mathrm{per}(A) = (-1)^{m-1} \sum_{U \subset \{2,\dots,m\}} \det(A_U) \cdot \det(A_{\bar{U}}),$$

where A_U is the principal submatrix of A induced by the rows and columns of A indexed by U, and $\bar{U}$ is the complement of U in $\{1, \dots, m\}$.

Proof. For $i, j, k, \dots \in \{1, \dots, m\}$ define $A_{i,j,k,\dots}$ to be the principal submatrix of A induced by removing the rows and columns indexed by $i, j, k, \dots$.

From Corollary 3.3

$$\begin{aligned}
\mathrm{per}(A) &= (-1)^m \left(\prod_{i=1}^m \frac{\partial}{\partial a_{ii}}\right) \det(A)^2 = (-1)^m \left(\prod_{i=2}^m \frac{\partial}{\partial a_{ii}}\right) \frac{\partial}{\partial a_{11}} \det(A)^2 \\
&= (-1)^m \left(\prod_{i=2}^m \frac{\partial}{\partial a_{ii}}\right) 2|A| \cdot |A_1| = (-1)^{m-1} \left(\prod_{i=3}^m \frac{\partial}{\partial a_{ii}}\right) \frac{\partial}{\partial a_{22}} |A| \cdot |A_1| \\
&= (-1)^{m-1} \left(\prod_{i=3}^m \frac{\partial}{\partial a_{ii}}\right) |A_2| \cdot |A_1| + |A| \cdot |A_{1,2}| \\
&= (-1)^{m-1} \left(\prod_{i=4}^m \frac{\partial}{\partial a_{ii}}\right) \frac{\partial}{\partial a_{33}} |A_2| \cdot |A_1| + |A| \cdot |A_{1,2}| \\
&= (-1)^{m-1} \left(\prod_{i=4}^m \frac{\partial}{\partial a_{ii}}\right) |A_{2,3}| \cdot |A_1| + |A_2| \cdot |A_{1,3}| + |A_3| \cdot |A_{1,2}| + |A| \cdot |A_{1,2,3}| \\
&= \dots = (-1)^{m-1} \sum_{U \subset \{2,\dots,m\}} |A_U| \cdot |A_{\bar{U}}|. \quad □
\end{aligned}$$

References

[1] R.A. Brualdi, H.J. Ryser, Combinatorial Matrix Theory, in: Encyclopedia of Mathematics and its Applications, vol. 39, Cambridge University Press, Cambridge, 1991.
[2] L.E. Dickson, History of the Theory of Numbers. Vol. 1: Divisibility and Primality, Chelsea Publishing Company, New York, 1966.
[3] D.G. Glynn, The modular counterparts of Cayley's hyperdeterminants, Bull. Aust. Math. Soc. 57 (1998) 479–492.
[4] R. Goodman, N.R. Wallach, Representations and Invariants of the Classical Groups, in: Encyclopedia of Mathematics and its Applications, vol. 68, Cambridge University Press, Cambridge, 1968.
[5] J.H. van Lint, R.M. Wilson, A Course in Combinatorics, Cambridge University Press, Cambridge, 1991.