ASYMPTOTIC BEHAVIOR OF THE EXPECTATION VALUE OF PERMANENT PRODUCTS

PAUL FEDERBUSH*

ABSTRACT. We would desire to have done the calculations of this paper in the measure on $n \times n$ matrices that weights uniformly all $0-1$ matrices with row and column sums equal to r, other matrices given weight zero. Instead we work with all matrices that are the sum of r independent uniformly weighted permutation matrices, with the hope that the computations we perform give the same result in this measure. We derive the result for limiting expectations

$$\lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_m(A)\mathrm{perm}_{m'}(A))) = \lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_m(A)))+ \lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_{m'}(A)))$$

Here r is fixed and m and m' are taken as each proportional to n.

So we work with a measure on matrices giving uniform weight to a sum of r independent permutation matrices. [1]. We have as a goal the formula

$$\lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_m(A)\,\mathrm{perm}_{m'}(A))) = \lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_m(A))) + \lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_{m'}(A))) \tag{1}$$

where r is fixed, and m and m' are each proportional to n. The result may seem more reasonable in light of Friedland's Asymptotic Matching Conjecture, now proven. [2, 3, 4]. Many readers will feel the computation of this paper can in a straightforward manner be made into a rigorous proof. But we do make an assumption later, the "statistical assumption", that though "obvious", should probably have a formal proof. We plan to present such a proof in a future paper.

We start with the computation of the expectation of $\mathrm{perm}_m(A)$ given in [5], which serves as a model for us. Their equation (4.3):

$$E(\mathrm{perm}_m A) = \frac{1}{(n!)^r} \binom{n}{m}^2 m! \sum_{m_1,\ldots,m_r \in \mathbb{Z}_+, m_1+\ldots m_r = m} \frac{m!(n-m_1)!\ldots(n-m_r)!}{m_1!\ldots m_r!} \tag{2}$$

We let $F(z)$ be given by

$$F(z) = z\ln z - z \tag{3}$$

We define the set S to be the region

$$\begin{aligned} &m_i \geq 0 \quad i = 1, \cdots, r-1 \\ &m_1 + \cdots + m_{r-1} \leq m \end{aligned} \tag{4}$$

and set

$$m_r = m - m_i - \cdots - m_{r-1} \tag{5}$$

*DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109-1043, *email:* PFED@UMICH.EDU

Date: July 21, 2014.

We define

$$R = -rF(n) + 2F(n) - 2F(n-m) - \sum_i F(m_i) + \sum_i F(n-m_i) \tag{6}$$

We now have as an approximation to (2) the formula

$$\int_S dm_1 \cdots dm_{r-1} e^R \tag{7}$$

The maximum of R inside the region of integration is achieved when all m_i are equal, so $m_i = \frac{m}{r}$. Let R_0 be this maximum value of R. Then

$$\lim_{n\to\infty} \frac{1}{n} \ln(E(\mathrm{perm}_m)) = \frac{R_0}{n} \tag{8}$$

With

$$m = pn \tag{9}$$

and so $m_i = \dfrac{pn}{r}$, $\dfrac{R_0}{n}$ becomes

$$-p\ln p + (2p-r)\ln r + 2(p-1)\ln(1-p) + (r-p)\ln(r-p) \tag{10}$$

We will play the same game with the expectation $E(\mathrm{perm}_m\, \mathrm{perm}_{m'})$.

We introduce some notation. We view the sum in (2) as over "$m - terms$." An m-term is an (unordered) set of m elements, each element having a location and a color. The color is an integer in the set $1, 2, \cdots, r$, and the location is an assignment of a row and a column in a universal $n \times n$ matrix. There is the restriction that two distinct elements in an m-term cannot have the same row or the same column in their locations. In (2) the expression

$$\binom{n}{m}^2 m! \tag{11}$$

sums over the locations, and the expression

$$\sum_{m_i} \frac{m!}{m_1! \cdots m_r!} \tag{12}$$

sums over the assignments of color.

In computing $E(\mathrm{perm}_m\, \mathrm{perm}_{m'})$ we will be summing over pairs consisting of an m-term and an m′-term. There is the additional restriction that if an element of the m-term is of the same color as an element of the m′-term, and if these two elements have the same row or column in their locations, then they have the same location (i.e. they have the same row and same column). Dealing with this restriction is the source of all the difficulties.

We construct an exact expression for the expectation of a product of permanents, a generalization of (2)

$$E(\mathrm{perm}_m\, \mathrm{perm}_{m'}) = M \circ A \circ E \circ B \circ C \circ D \circ T \tag{13}$$

Each module in the right side of (13), except the last, includes sums. The sums are to be performed from right to left, the same order as in integration.

The module M sums over the m-terms in perm_m

$$M = \sum_{m_i} \frac{1}{(n!)^r} \binom{n}{m}^2 m! \frac{m!}{m_1! \cdots m_r!} \tag{14}$$

m_i is the number of elements in a given m-term of color i. Equation (5) holds. We have included the normalization factor $\frac{1}{(n!)^r}$ in M. Specification of ranges of sums is postponed to later. Note that we thus sum over the m'-terms in $\mathrm{perm}_{m'}$ in a "background" of a fixed m-term, later summed over by M. A, E, B, C, and D sum over these m$'$-terms.

With the fixed background m-term (laid down by M) given, we divide the elements of any m$'$-term into five classes.

class 1: An element is in class 1 if the row and column of its location are disjoint from all the rows and columns associated to elements in the fixed m-term. class 1 is summed over by A.

class 2: An element is in class 2 if it agrees with some element in the fixed m-term in both color and location. class 2 is summed over by E.

class 3: An element is in class 3 if the row of its location agrees with the row assignment of some element in the fixed m-terms, but its column assignment is disjoint from that of all elements in the fixed m-term. class 3 is summed over by B.

class 4: An element is in class 4 if the column of its location agrees with the column assignment of some element in the fixed m-term, but its row assignment is disjoint from that of all elements in the fixed m-term. class 4 is summed over by C.

class 5: An element is in class 5 if its row assignment agrees with the row assignment of some element in the fixed m-term, and likewise its column assignment agrees with the column assignment of some element in the fixed m-term, but it does not agree in both color and location with any element in the m-term. class 5 is summed over by D.

For the m$'$-term being summed up let a be the number of its elements in class 1, e the number in class 2, b the number in class 3, c the number in class 4, and d the number in class 5. The classes are disjoint and cover all cases, so one has

$$a + e + b + c + d = m' \tag{15}$$

For each letter adding a single subscript specifies the number of a given color. So b_i is the number of elements in class 3 of color i. One has

$$a = \sum_i a_i \tag{16}$$

$$e = \sum_i e_i \tag{17}$$

$$b = \sum_i b_i \tag{18}$$

$$c = \sum_i c_i \tag{19}$$

$$d = \sum_i d_i \tag{20}$$

The module T sums over all possibilities for the r permutation matrices not determined by the selection of the given m-term and m$'$-term.

$$T = \prod_i (n - m_i - a_i - b_i - c_i - d_i)! \tag{21}$$

This is the generalization of the

$$\prod_i (n - m_i)! \tag{22}$$

factor in (2).

The <u>module A</u> sums over all possibilities for the number of elements in the m′-term of class 1.

$$A = \sum_{a_i} \binom{n-m}{a}^2 a! \left(\frac{a!}{a_1! \cdots a_r!} \right) \tag{23}$$

<u>Module E</u> sums over all possibilities for the number of elements in the m′-term of class 2.

$$E = \sum_{e_i} \prod_i \binom{m_i}{e_i} \tag{24}$$

We now consider <u>module B</u> which sums over the number of elements in the m′-term of class 3. We have non-negative integers b_{ik} with $i, k \in \{1, 2, \cdots, r\}$ and $i \neq k$.

$$b_i = \sum_{k \neq i} b_{ik} \tag{25}$$

We also define

$$\widetilde{b_i} = \sum_{k \neq i} b_{ki} \tag{26}$$

b_{ik} is the number of elements in the m′-term that are of class 3 and color i and which share a row with an element of the m-term of color k.

$$B = \sum_{b_{ik}} \left[\binom{n-m-a}{b} \frac{b!}{b_1! \cdots b_r!} \prod_i \frac{b_i!}{\prod_{k \neq i} b_{ik}!} \right] \cdot \left[\prod_i \binom{m_i - e_i}{\widetilde{b_i}} \frac{\widetilde{b_i}}{\prod_{k \neq i} b_{ki}!} \right] \cdot \left[\prod_i \prod_{k \neq i} b_{ik}! \right] \tag{27}$$

The first bracket selects the columns to be assigned to the elements of each set b_{ik}, the second bracket selects the rows to be assigned to the elements of each set b_{ik}, and the third bracket counts the ways of grouping the rows and columns.

<u>Module C</u> is constructed as module B with rows and columns interchanged. c_i and c_{ik} replace b_i and b_{ik}.

$$c_i = \sum_{k \neq i} c_{ik} \tag{28}$$

$$\widetilde{c_i} = \sum k \neq i c_{ki} \tag{29}$$

$$C = \sum_{c_{ik}} \left[\binom{n-m-a}{c} \frac{c!}{c_1! \cdots c_r!} \prod_i \frac{c_i!}{\prod_{k \neq i} c_{ik}!} \right] \cdot \left[\prod_i \binom{m_i - e_i}{\widetilde{c_i}} \frac{\widetilde{c_i}}{\prod_{k \neq i} c_{ki}!} \right] \cdot \left[\prod_i \prod_{k \neq i} c_{ik}! \right] \tag{30}$$

<u>Module D</u> sums over the possibilities for the number of elements in the m′-term of class 5. We have non-negative integers l_{ik} and r_{ik} with $i, k \in \{1, \cdots, r\}, i \neq k$.

$$d_i = \sum_{k \neq i} l_{ik} = \sum_{k \neq i} r_{ik} \tag{31}$$

$$\widetilde{l_i} = \sum_{k \neq i} l_{ki} \tag{32}$$

$$\widetilde{r_i} = \sum_{k \neq i} r_{ki} \tag{33}$$

l_{ik} is the number of elements in the m′-term of class 5 and color i and which share a row with an element of the m-term of color k. r_{ik} has a similar definition with column replacing row.

$$D = \sum_{r_i k} \sum_{l_i k} \prod_i \left[\binom{m_i - e_i - \widetilde{b_i}}{\widetilde{l_i}} \cdot \frac{\widetilde{l_i}!}{\prod_{k \neq i} l_{ki}!} \right] \cdot \left[\prod_i \binom{m_i - e_i - \widetilde{c_i}}{\widetilde{r_i}} \cdot \frac{\widetilde{r_i}!}{\prod_{k \neq i} r_{ki}!} \right] \cdot [d_i!] \tag{34}$$

We turn to the discussion of the ranges of the sums in (13). The primary variables, non-negative integers, are m_i, a_i, e_i, b_{ik}, c_{ik}, l_{ik}, and r_{ik}. From these are defined secondary variable a, e, b, c, d, b_i, $\widetilde{b_i}$, c_i, $\widetilde{c_i}$, d_i, $\widetilde{l_i}$, $\widetilde{r_i}$, by equations (16), (17), (18), (19), (20), (25), (26), (28), (29), (31), (32), (33) respectively. Restrictions on the independence of the basic variables are given by (5), (15), and (31). In addition to these three restrictions we find the following inequalities:

(a) From the consideration of (21) we find

$$m_i + a_i + b_i + c_i + d_i \leq n, \text{ each } i \tag{35}$$

(b) From the consideration of (23) we find

$$a \leq n - m \tag{36}$$

(c) From the consideration of (24) we find

$$e_i \leq m_i, \text{ each } i \tag{37}$$

(d) From the consideration of (27) we find

$$a + b + m \leq n \tag{38}$$

and

$$e_i + \widetilde{b_i} \leq m_i, \text{ each } i \tag{39}$$

(e) From the consideration of (30) we find

$$a + c + m \leq n \tag{40}$$

and

$$e_i + \widetilde{c_i} \leq m_i, \text{ each } i \tag{41}$$

(f) From the consideration of (34) we find

$$e_i + \widetilde{b_i} + \widetilde{l_i} \leq m_i, \text{ each } i \tag{42}$$

and

$$e_i + \widetilde{c_i} + \widetilde{r_i} \leq m_i, \text{ each } i \tag{43}$$

With the definition of the modules, M through T, and these restrictions on the sum over the primary variables, equation (13) provides an exact expression for the expectation of the product of permanents.

Referring back to the developments surrounding equation (7) we write the exact expression for the expectation in the form

$$\sum e^S \tag{44}$$

We consider S as a function of the primary variables. Then the "statistical assumption" says that if a, e, b, c, d, are fixed at some large values, that is on the scale of n, then with these restrictions, the largest value of S is obtained when

$$m_i = m_j, \text{ all } i,j \tag{45}$$

$$a_i = a_j, \text{ all } i,j \tag{46}$$

$$e_i = e_j, \text{ all } i,j \tag{47}$$

$$b_{ij} = b_{kl}, \text{ all } i,j,k,l \tag{48}$$

$$c_{ij} = c_{kl}, \text{ all } i,j,k,l \tag{49}$$

$$l_{ij} = l_{kl}, \text{ all } i,j,k,l \tag{50}$$

$$r_{ij} = r_{kl}, \text{ all } i,j,k,l \tag{51}$$

$$l_{ij} = r_{ij}, \text{ all } i,j \tag{52}$$

Relations (45) through (52) holding on the scale of 1.

These equalities we argue hold by statistical considerations of how the colors distribute themselves once the structure of the locations is specified.

We now set

$$m_i = \frac{m}{r} \tag{53}$$

$$a_i = \frac{a}{r} \tag{54}$$

$$e_i = \frac{e}{r} \tag{55}$$

$$b_{ij} = \frac{b}{r(r-1)} \tag{56}$$

$$c_{ij} = \frac{c}{r(r-1)} \tag{57}$$

$$l_{ij} = r_{ij} = \frac{d}{r(r-1)} \tag{58}$$

We also set $b = c$ for simplicity as will turn out true if we let them remain independent through the following discussion. We use the approximation (3) for $\ln(z!)$ and write S using (53) through (58) with $b = c$.

$$S \cong LM + LA + LE + 2LB + LD + LT \tag{59}$$

with

$$LM = -rF(n) + 2F(n) - 2F(n-m) - rF\left(\frac{m}{r}\right) \tag{60}$$

from (14), and

$$LA = 2F(n-m) - 2F(n-m-a) - rF(\frac{a}{r}) \tag{61}$$

from (23), and

$$LE = rF(\frac{m}{r}) - rF(\frac{m}{r} - \frac{e}{r}) - rF\left(\frac{e}{r}\right) \tag{62}$$

from (24), and

$$LB = F(n-m-a) - F(n-m-a-b) - r(r-1)F\left(\frac{b}{r(r-a)}\right) + rF\left(\frac{m}{r} - \frac{e}{r}\right) - rF\left(\frac{m}{r} - \frac{e}{r} - \frac{b}{r}\right) \tag{63}$$

from (27), and

$$LD = 2rF\left(\frac{m-e-b}{r}\right) - 2r(r-1)F\left(\frac{d}{r(r-1)}\right) + rF\left(\frac{d}{r}\right) - 2rF\left(\frac{m-e-b-d}{r}\right) \quad (64)$$

from (34), and

$$LT = rF\left(n - \frac{m+a+2b+d}{r}\right) \quad (65)$$

from (21).

Using Lagrange multipliers to impose equation (15) we write

$$SS = S - l\cdot(a+e+2b+d) \quad (66)$$

and attempt to solve the five following equations together.

$$\frac{\partial SS}{\partial a} = 0 \text{ or } (n-m-a-b)^2 - L\left(\frac{a}{r}\right)\cdot\left(n - \frac{m+a+2b+d}{r}\right) = 0 \quad (67)$$

$$\frac{\partial SS}{\partial b} = 0 \text{ or } (n-m-a-b)\cdot\left(\frac{m-e-b-d}{r}\right) - L\left(\frac{b}{r(r-1)}\right)\cdot\left(n - \frac{m+a+2b+d}{r}\right) = 0 \quad (68)$$

$$\frac{\partial SS}{\partial d} = 0 \text{ or } \left(\frac{(m-e-b-d)}{r}\right)^2 - Ld\frac{1}{(r(r-1)^2)}\cdot\left(n - \frac{m+a+2b+d}{r}\right) = 0 \quad (69)$$

$$\frac{\partial SS}{\partial e} = 0 \text{ or } \left(\frac{(m-e-b-d)}{r}\right)^2 - L\left(\frac{m}{r} - \frac{e}{r}\right)\cdot\left(\frac{e}{r}\right) = 0 \quad (70)$$

$$a+e+2b+d = m' \quad (71)$$

Amazingly enough Maple found the following analytic solution to this problem. We write $m = pn$ and $m' = qn$ and then have

$$a = nq(1-p)^2\frac{r}{(r-p)} \quad (72)$$

$$b = n(1-p)(r-1)\frac{pq}{(r-p)} \quad (73)$$

$$d = n(r-1)^2p^2\frac{q}{(r(r-p))} \quad (74)$$

$$e = n\frac{pq}{r} \quad (75)$$

$$L = (1-q)^2\frac{r^2}{q(r-q)} \quad (76)$$

Substituting these values into S we find, again amazingly

$$S = n\Big[\Big(-p\ln p + (2p-r)\ln r + 2(p-1)\ln(1-p) + (r-p)\ln(r-p)\Big) + \Big(-q\ln q + (2q-r)\ln r + 2(q-1)\ln(1-q) + (r-q)\ln(r-q)\Big)\Big] \quad (77)$$

References

[1] H.S. Wilf, On the permanent of a doubly stochastic matrix, Canad. J. Math. 18 (1966) 758-761

[2] S. Friedland, E. Krop. Per H. Lundow, K. Markström; Validations of the Asymptotic Matching Conjectures, arXiv:math/0603001, 2006

[3] L. Gurvits: *Unleashing the power of Schrijver's permanental inequality with the help of the Bethe Approximations*, arXiv 1106.2844v11

[4] P. Csikvari, Lower Matching Conjecture, and a New Proof of Schrijver's and Gurvits's Theorems, arXiv:1406.0766 [math.co]

[5] S. Friedland, E. Krop and K. Markström: *On the Number of Matchings in Regular Graphs*, The Electronic Journal of Combinatorics, **15** (2008), R110, pp. 1-28.