The n-card problem, stochastic matrices, and the Extreme Principle

Justin H.C. Chan Jonathan Jedwab*

Department of Mathematics
Simon Fraser University
Burnaby, BC, Canada

{jhc34,jed}@sfu.ca

Submitted: Jan 16, 2012; Accepted: Jun 17, 2012; Published: Jun 28, 2012
Mathematics Subject Classifications: 05A05

Abstract

The n-card problem is to determine the minimal intervals $[u, v]$ such that for every $n \times n$ stochastic matrix A there is an $n \times n$ permutation matrix P (depending on A) such that $\mathrm{tr}(PA) \in [u, v]$. This problem is closely related to classical mathematical problems from industry and management, including the linear assignment problem and the travelling salesman problem. The minimal intervals for the n-card problem are known only for $n \leqslant 4$.

We introduce a new method of analysis for the n-card problem that makes repeated use of the Extreme Principle. We use this method to answer a question posed by Sands (2011), by showing that $[1, 2]$ is a solution to the n-card problem for all $n \geqslant 2$. We also show that each closed interval of length $\frac{n}{n-1}$ contained in $[0, 2)$ is a solution to the n-card problem for all $n \geqslant 2$.

Keywords: stochastic matrix; permutation matrix; transversal sum; trace; Extreme Principle; n-card problem

1 Introduction

Let $n \geqslant 2$ be an integer. An $n \times n$ *stochastic matrix* is an $n \times n$ matrix (a_{ij}) of non-negative real numbers, each of whose row sums is 1. A *transversal sum* of (a_{ij}) is a sum of the form $\sum_{i=1}^{n} a_{\sigma(i),i}$, for some permutation σ of $\{1, 2, \ldots, n\}$. A *solution to the n-card problem* is an interval $[u, v]$ such that every $n \times n$ stochastic matrix contains at least one transversal sum in $[u, v]$. Equivalently, a solution to the n-card problem is an interval $[u, v]$ such that

*Supported by NSERC.

for every $n \times n$ stochastic matrix A there is an $n \times n$ permutation matrix P such that $\operatorname{tr}(PA) \in [u, v]$. We wish to determine the minimal solutions to the n-card problem for each n, namely those solutions $[u, v]$ for which no proper subinterval of $[u, v]$ is a solution.

The n-card problem is closely related to well-known mathematical problems with industrial and management applications, involving the possible values of $\operatorname{tr}(PA)$ for a permutation matrix P and a fixed square matrix A. In particular, the linear assignment problem, "one of the most famous problems in linear programming and in combinatorial optimization [BDM09]," is to minimize $\operatorname{tr}(PA)$ over all permutation matrices P (see [BDM09, Chapter 4] for a detailed historical account of the development of algorithms for its solution, and an equivalent formulation in terms of weighted bipartite matchings); the travelling salesman problem is the special case in which the permutation corresponding to P is cyclic [Flo56].

The terminology of the n-card problem reflects its original formulation [San01], [LS05] involving a set of n cards, each containing n non-negative real numbers written in a row and summing to 1, with the transversal sum representing the diagonal sum formed when the cards are placed one below the other according to some permutation. In 2001, Sands [San01] asked for a proof that $\left[\frac{1}{2}, \frac{3}{2}\right]$ is a solution to the 3-card problem. Lenza and Sands [LS05] introduced the generalization to the n-card problem in 2005.

The interval $[0, 1]$ is a minimal solution to the n-card problem for all $n \geqslant 2$ [San11, Lemma 3], but it is not the only minimal solution. Indeed, it is easily checked by hand that the minimal solutions to the 2-card problem are

$$[0, 1], \quad [1, 2]. \tag{1}$$

The minimal solutions to the 3-card problem are [LS05]

$$[0, 1], \quad \left[\tfrac{1}{2}, \tfrac{3}{2}\right], \quad [1, 2], \tag{2}$$

and the minimal solutions to the 4-card problem are [LS05], [San11]

$$[0, 1], \quad \left[\tfrac{1}{3}, \tfrac{4}{3}\right], \quad \left[\tfrac{2}{3}, \tfrac{5}{3}\right], \quad [1, 2]. \tag{3}$$

The method of [LS05] and [San11] is to establish a particular interval $[u, v]$ as a solution to the n-card problem (for $n = 3$ or 4), by using intersecting permutations to show that the number of transversal sums greater than v, plus the number of transversal sums less than u, is always less than $n!$. The method has two drawbacks: it relies on a laborious case analysis for $n = 4$, and does not extend to $n \geqslant 5$ [LS05, p.6].

In this paper we introduce a new method for analysing the n-card problem that makes repeated use of the Extreme Principle [Zei07]. We believe that this method could shed light on other problems involving $\operatorname{tr}(PA)$, where P is a permutation matrix and A is a fixed square matrix. The Extreme Principle directs attention to the "largest" and "smallest" elements of a problem. In the present context, we assume for a contradiction that no transversal sum of an $n \times n$ stochastic matrix lies in some interval $[u, v]$, and then consider the smallest transversal sum d greater than v. Then, if a transversal sum is less than d, it must be less than u. We seek such transversal sums, involving exactly $n - 2$ of the original summands of d, in order to reach a contradiction. We thereby obtain strong new restrictions for all $n \geqslant 5$. In particular, we solve Problem 5 of [San11] as follows.

Theorem 1. *For all $n \geqslant 2$, the interval $[1,2]$ is a solution to the n-card problem.*

We also prove the following result.

Theorem 2. *For all $n \geqslant 2$, each closed interval of length $\frac{n}{n-1}$ contained in $[0,2)$ is a solution to the n-card problem.*

The length $\frac{n}{n-1}$ in Theorem 2 is the smallest possible for a general interval, because the $n \times n$ stochastic matrix having one row $\left(0, \frac{1}{n-1}, \frac{1}{n-1}, \ldots, \frac{1}{n-1}\right)$ and $n-1$ rows $(1,0,0,\ldots,0)$ has transversal sums of 0 and $\frac{n}{n-1}$ only.

On the other hand, the known complete set of minimal solutions (2) for $n = 3$ and (3) for $n = 4$ shows that the interval length in Theorem 2 can be reduced to 1 for specific intervals. By reference to particular $n \times n$ stochastic matrices, Sands [San11] showed that every solution to the n-card problem for $n \geqslant 2$ must contain a length 1 interval $\left[\frac{k}{n-1}, 1 + \frac{k}{n-1}\right]$ for some $k \in \{0, 1, \ldots, n-1\}$. Problem 3 of [San11] asks whether each such length 1 interval is itself a solution to the n-card problem, which would imply that the complete set of minimal solutions to the n-card problem comprises these n intervals. This question remains open for $n > 4$.

2 The interval [1,2]

In this section we prove Theorem 1. We firstly establish some preliminary lemmas.

Lemma 3. *Let (a_{ij}) be an $n \times n$ stochastic matrix, all of whose transversal sums lie outside an interval $[u,v]$ containing 1. Then (a_{ij}) has at least one transversal sum less than u, and at least one transversal sum greater than v.*

Proof. Each entry of (a_{ij}) is contained in exactly $(n-1)!$ transversal sums, so the mean of all transversal sums is $(n-1)!(\sum_{i,j} a_{ij})/n! = 1$. Therefore at least one transversal sum is at most 1, and so by assumption less than u. Similarly, at least one transversal sum is greater than v. □

An immediate consequence of Lemma 3 is that, as noted earlier, $[0,1]$ is a solution to the n-card problem for all $n \geqslant 2$.

The rows and columns of an $n \times n$ stochastic matrix can be permuted without changing the set of its $n!$ transversal sums. Our method relies on examining the effect of transposing two rows of an $n \times n$ stochastic matrix, and thereby bounding the matrix entries. We now show that if an $n \times n$ stochastic matrix has a sufficiently large diagonal sum then there must be a transposition of two rows that decreases this diagonal sum. We prove this result for the following slightly more general case of an $n \times n$ substochastic matrix (each of whose row sums is at most 1).

Lemma 4. *Let (a_{ij}) be an $n \times n$ substochastic matrix. Suppose (a_{ij}) has diagonal sum greater than* 1*. Then*

$$a_{kk} + a_{\ell\ell} > a_{k\ell} + a_{\ell k} \quad \textit{for some } k, \ell.$$

Proof. Suppose, for a contradiction, that $a_{ii}+a_{jj} \leqslant a_{ij}+a_{ji}$ for all i, j. Sum this inequality over all i, j to obtain $2n \sum_i a_{ii} \leqslant 2 \sum_{i,j} a_{ij} \leqslant 2n$, since by assumption the row sums of (a_{ij}) are each at most 1. This implies that the diagonal sum satisfies $\sum_i a_{ii} \leqslant 1$, giving the required contradiction. □

We next give conditions under which the sum of two diagonal entries of an $n \times n$ stochastic matrix can be bounded from below.

Lemma 5. *Let (a_{ij}) be an $n \times n$ stochastic matrix with diagonal sum d, and suppose all transversal sums of (a_{ij}) lie outside the interval $[u, d)$. Then, for all i, j,*

$$a_{ii} + a_{jj} > a_{ij} + a_{ji} \quad \textit{implies} \quad a_{ii} + a_{jj} > d - u.$$

Proof. Suppose $a_{ii} + a_{jj} > a_{ij} + a_{ji}$. Then the positive quantity $a_{ii} + a_{jj} - a_{ij} - a_{ji}$ is the decrease in the diagonal sum caused by transposing rows i and j of the matrix, and so by assumption is greater than $d-u$. We therefore have $a_{ii}+a_{jj} \geqslant a_{ii}+a_{jj}-a_{ij}-a_{ji} > d-u$. □

Define an $n \times n$ stochastic matrix (a_{ij}) to be *diagonally ordered* if its diagonal entries are in non-increasing order:

$$a_{11} \geqslant a_{22} \geqslant \ldots \geqslant a_{nn}.$$

We are now ready to prove Theorem 1.

Proof of Theorem 1. We know from (1) and (2) that the result holds for $n = 2$ and 3, so we may take $n \geqslant 4$. Suppose, for a contradiction, that (a_{ij}) is an $n \times n$ stochastic matrix whose transversal sums all lie outside the interval $[1, 2]$. Then by Lemma 3, (a_{ij}) has a transversal sum greater than 2 and a transversal sum less than 1. Let $2 + \epsilon$ be the smallest transversal sum greater than 2, and reorder the rows and columns of (a_{ij}) so that the summands of this transversal sum occur on the matrix diagonal and so that the matrix is diagonally ordered. By Lemma 5 with $d = 2 + \epsilon$ and $u = 1$,

$$a_{ii} + a_{jj} > a_{ij} + a_{ji} \quad \text{implies} \quad a_{ii} + a_{jj} > 1 + \epsilon. \tag{4}$$

Now the $(n-1) \times (n-1)$ submatrix of (a_{ij}) formed by deleting the first row and column has diagonal sum $2 + \epsilon - a_{11} > 1$. Apply Lemma 4 to this submatrix to show that, for some distinct $k > 1$ and $\ell > 1$,

$$a_{kk} + a_{\ell\ell} > a_{k\ell} + a_{\ell k}.$$

We therefore have $a_{kk} + a_{\ell\ell} > 1 + \epsilon$ by (4), and so

$$a_{22} + a_{33} > 1 + \epsilon \tag{5}$$

since the matrix is diagonally ordered and k, ℓ are distinct. Since the diagonal sum of (a_{ij}) is $2+\epsilon$, we have

$$a_{ii} + a_{11} \leqslant 2 + \epsilon - a_{22} - a_{33} \quad \text{for all } i > 3,$$

and therefore $a_{ii} + a_{11} < 1$ for all $i > 3$, by (5). Then, since the matrix is diagonally ordered,

$$a_{ii} + a_{jj} < 1 \quad \text{for all } i, j \text{ with } i > 3,$$

which in turn implies by (4) that

$$a_{ii} + a_{jj} \leqslant a_{ij} + a_{ji} \quad \text{for all } i, j \text{ with } i > 3. \tag{6}$$

We complete the proof by showing that (5) and (6) force the sum of the entries of (a_{ij}) to be too large. We have

$$\begin{aligned}\sum_{i,j} a_{ij} &\geqslant \sum_{i \leqslant 3} a_{ii} + \sum_{i>3} a_{ii} + \sum_{i>3,\, j \leqslant 3} (a_{ij} + a_{ji}) \\ &\geqslant (n-2) \sum_{i \leqslant 3} a_{ii} + 4 \sum_{i>3} a_{ii}\end{aligned}$$

by substitution from (6). Therefore

$$\begin{aligned}\sum_{i,j} a_{ij} &\geqslant (n-2) \sum_{i \leqslant 3} a_{ii} + 2 \sum_{i>3} a_{ii} \\ &= (n-4) \sum_{i \leqslant 3} a_{ii} + 2 \sum_{i} a_{ii} \\ &\geqslant (n-4)(1+\epsilon) + 2(2+\epsilon)\end{aligned}$$

by (5), using $n \geqslant 4$. Therefore $\sum_{i,j} a_{ij} > n$, which is a contradiction because each row sum of (a_{ij}) is 1. □

3 Intervals of length $\frac{n}{n-1}$

In this section we prove Theorem 2.

Proposition 6. *Let $n \geqslant 4$ and let (a_{ij}) be a diagonally ordered $n \times n$ stochastic matrix. Suppose the diagonal sum d of (a_{ij}) satisfies $d \in (1, 2]$. Then (a_{ij}) has a transversal sum lying in the interval $[d - \frac{n}{n-1}, d)$.*

Proof. Suppose, for a contradiction, that no transversal sum of (a_{ij}) lies in the interval $[d - \frac{n}{n-1}, d)$. Then by Lemma 5 with $u = d - \frac{n}{n-1}$,

$$a_{ii} + a_{jj} > a_{ij} + a_{ji} \quad \text{implies} \quad a_{ii} + a_{jj} > \frac{n}{n-1}. \tag{7}$$

Since $d > 1$, Lemma 4 gives

$$a_{kk} + a_{\ell\ell} > a_{k\ell} + a_{\ell k} \quad \text{for some distinct } k, \ell,$$

and it follows from (7) that $a_{kk} + a_{\ell\ell} > \frac{n}{n-1}$. Since the matrix is diagonally ordered and k, ℓ are distinct, this implies

$$a_{11} + a_{22} > \frac{n}{n-1} \tag{8}$$

and

$$a_{11} > \frac{1}{2} \cdot \frac{n}{n-1}. \tag{9}$$

We now claim that

$$a_{ii} + a_{jj} \leqslant \frac{n}{n-1} \quad \text{for all distinct } i > 1 \text{ and } j > 1. \tag{10}$$

Suppose otherwise, for a contradiction, so that $a_{rr} + a_{ss} > \frac{n}{n-1}$ for some distinct $r > 1$ and $s > 1$. Since the matrix is diagonally ordered, this gives

$$a_{22} + a_{33} > \frac{n}{n-1}. \tag{11}$$

Therefore, for all $i > 3$, we have $a_{ii} + a_{11} \leqslant d - a_{22} - a_{33} < \frac{n-2}{n-1}$ because $d \leqslant 2$. Since the matrix is diagonally ordered, we then have

$$a_{ii} + a_{jj} < \frac{n-2}{n-1} \quad \text{for all } i, j \text{ with } i > 3,$$

which by (7) implies

$$a_{ii} + a_{jj} \leqslant a_{ij} + a_{ji} \quad \text{for all } i, j \text{ with } i > 3. \tag{12}$$

Write

$$\begin{aligned} \sum_{i,j} a_{ij} &\geqslant \sum_{i \leqslant 3} a_{ii} + \sum_{i>3, j \leqslant 3} (a_{ij} + a_{ji}) \\ &\geqslant (n-2) \sum_{i \leqslant 3} a_{ii} + 3 \sum_{i>3} a_{ii} \end{aligned}$$

by substitution from (12), so that

$$\begin{aligned} \sum_{i,j} a_{ij} &\geqslant (n-2) \sum_{i \leqslant 3} a_{ii} \\ &> (n-2) \cdot \frac{3}{2} \cdot \frac{n}{n-1} \end{aligned}$$

from (9) and (11). Since $\sum_{ij} a_{ij} = n$ and $n \geqslant 4$, this is a contradiction and proves the claim (10).

It then follows from (7) that

$$a_{ii} + a_{jj} \leqslant a_{ij} + a_{ji} \quad \text{for all distinct } i > 1 \text{ and } j > 1.$$

Summing over all i, j satisfying $1 < i < j$, we find that

$$(n-2) \sum_{i>1} a_{ii} \leqslant \sum_{1<i<j} (a_{ij} + a_{ji}). \tag{13}$$

Now let m be the largest integer i such that $a_{11} + a_{ii} > \frac{n}{n-1}$. Note that $m \geqslant 2$, by (8). By (7) we have $a_{11} + a_{ii} \leqslant a_{1i} + a_{i1}$ for $i > m$, so that

$$a_{11} \leqslant a_{1i} + a_{i1} \quad \text{for } i > m. \tag{14}$$

We now show that (13) and (14) force the entries of (a_{ij}) to be too large. We have

$$\begin{aligned}\sum_{i,j} a_{ij} &\geqslant a_{11} + \sum_{i>1} a_{ii} + \sum_{i>m} (a_{1i} + a_{i1}) + \sum_{1<i<j} (a_{ij} + a_{ji}) \\ &\geqslant (n-m+1)a_{11} + (n-1) \sum_{i>1} a_{ii}\end{aligned}$$

by substitution from (13) and (14). Therefore

$$\begin{aligned}\sum_{i,j} a_{ij} &\geqslant (n-m+1)a_{11} + (n-1) \sum_{1<i\leqslant m} a_{ii} \\ &> (n-m+1)a_{11} + (n-1)(m-1)\left(\frac{n}{n-1} - a_{11}\right)\end{aligned}$$

by definition of m and the diagonal ordering of (a_{ij}), and so

$$\sum_{i,j} a_{ij} > n\big(m - 1 - (m-2)a_{11}\big).$$

Since $m \geqslant 2$ and $a_{11} \leqslant 1$, this implies the contradiction $\sum_{i,j} a_{ij} > n$ and so completes the proof. □

We now combine Proposition 6 with Theorem 1 to prove Theorem 2.

Proof of Theorem 2. We know from (1) and (2) that the result holds for $n = 2$ and 3, so we may take $n \geqslant 4$. Suppose, for a contradiction, that (a_{ij}) is an $n \times n$ stochastic matrix whose transversal sums all lie outside the interval $\big[u, u + \frac{n}{n-1}\big]$ for some $u \in \big[0, \frac{n-2}{n-1}\big)$. Since this interval contains 1, by Theorem 1 the matrix (a_{ij}) therefore has a transversal sum in the interval $\big(u + \frac{n}{n-1}, 2\big]$. Let d be the smallest such transversal sum. Reorder the rows and columns of (a_{ij}) so that the summands of this transversal sum occur on the matrix diagonal and so that the matrix is diagonally ordered. Then by Proposition 6, (a_{ij}) has a transversal sum lying in the interval $\big[d - \frac{n}{n-1}, d\big)$. By choice of d, this gives the required contradiction. □

Acknowledgements

The authors thank the organisers of the February 2011 *14th Coast Combinatorics Conference*, University of Victoria, BC, where they first learned of the n-card problem, and Prof. Bill Sands for kindly supplying a preprint of the paper [San11].

References

[BDM09] R. Burkard, M. Dell'Amico, and S. Martello. Assignment Problems. SIAM, Philadelphia, 2009.

[Flo56] M. M. Flood. The travelling-salesman problem. *Operations Research*, **4**: 61–75, 1956.

[LS05] E. Lenza and B. Sands. A new proof and extension of problem 2620. *Crux Mathematicorum with Mathematical Mayhem*, **31**:319–326, 2005.

[San01] B. Sands. Problem 2620. *Crux Mathematicorum with Mathematical Mayhem*, **27**: 138, 2001. Solution due to M. Bataille in **28**: 127–128, 2002.

[San11] B. Sands. Cards, permutations, and sums. *Contributions to Discrete Mathematics*, **6**: 1–19, 2011.

[Zei07] P. Zeitz. *The Art and Craft of Problem Solving*, 2nd edition, Wiley, 2007.