Partitio et Emergo

A Note on the Computation of Eigenvalues of a Shuffle Matrix

Frans Schalekamp
Anke van Zuylen

Abstract

The transition matrix of the Markov chain describing card shuffling ("shuffle matrix") is studied. The authors propose a method to compute (a subset of) the eigenvalues of a shuffle matrix, which is a generalization of a method proposed by Doner and Uppuluri [2]. The method works by defining Markov chains with smaller state spaces than the original state space, the transition matrices of which have eigenvalues which are a subset of the eigenvalues of the shuffle matrix.

corresponding author:
Frans Schalekamp
Cornell University
Operations Research & Industrial Engineering
257 Rhodes Hall
Ithaca NY 14853
fms9@cornell.edu

Anke van Zuylen
anke@orie.cornell.edu

*This note is the result of a master's thesis written while both authors were at the Vrije University, Amsterdam, The Netherlands, with Professor Henk Tijms as advisor.

1 Introduction

Shuffling has been an object of mathematical study for decades. It is easily seen, that card shuffling is a Markov chain: the result of a shuffle only depends on the order of the deck right before the shuffle is performed, and the shuffle method employed. The state space of this Markov chain, however, is gigantic, namely the number of permutations of the number of cards used.

This Markov chain's transition matrix has a special structure, which can be exploited for computing eigenvalues. In this note we give a generalization of a method proposed by Doner and Uppuluri [2], which exploits this structure. This generalization allows the computation of more eigenvalues than the original method. This raises the question whether all eigenvalues can be found this way, a question which we have not been able to answer.

Following Doner and Uppuluri [2], we study a "projection" of the shuffling process: the state space of this projection is obtained by partitioning the original state space. Note that this new process is not necessarily Markovian. Doner and Uppuluri [2] show that, if this new process is Markovian, the eigenvalues of the transition matrix of this Markov chain are a subset of the eigenvalues of the transition matrix of the original Markov chain. They show that for a very particular partition of the state space, the process is indeed Markovian, and moreover, that the transition matrix of this process is centrosymmetric, allowing a further speedup of the computation of the eigenvalues following Collar [1].

We will show that there are *multiple* ways of partitioning the state space, such that the random process defined on the partitions is Markovian, and the transition matrices corresponding to these Markov chains are centrosymmetric.

In our exploration, we stumbled upon a property of cyclic subgroups, which, to our knowledge, was not documented, yet.

2 Partitioning the state space of a shuffling process

For ease of notation, we will assume henceforth that the cards are numbered $1, \ldots, n$.

Notation We will write $\pi = (\pi^{(1)}\ \pi^{(2)}\ \cdots\ \pi^{(n)})$ for a permutation of length n, where $\pi^{(j)}$ is the number on the card in position j. We denote the algebraic inverse of π by π^{-1}. For the set of all permutations of length n, we will use the usual notation $\mathcal{S}_n$.

We define a shuffling method as a probability distribution on the operations that can be performed on a deck of cards. Note that the operations that can be performed can also be denoted by permutations of length n (one could also view these as the resulting order of the deck, if the order of the deck before shuffling was $(1\ 2\ 3\ \ldots\ n)$). We will thus use permutations to denote both the ordering of a deck, as well as the rearranging action.

Example of notation A deck in order (1 4 3 2) is shuffled using action (2 4 1 3). This results in the order (1 4 3 2)(2 4 1 3) = (4 2 1 3).

More notation Henceforth, we will assume that the state space is ordered lexicographically. Denote the permutations in lexicographical order as $\pi_1, \pi_2, \ldots, \pi_{n!}$.

Let M be the transition matrix of the shuffling process. Note that $M_{ij} \stackrel{\text{def}}{=} \mathbb{P}(\pi_i \to \pi_j) = \mathbb{P}(\text{action } (\pi_i^{-1}\pi_j)) = \mathbb{P}(\pi_1 \to \pi_i^{-1}\pi_j)$, using obvious notation. Note that there are at most $n!$ distinct entries in the transition matrix. Moreover, these entries are repeated in each row of the matrix (in a different order), as well as in each column.

A small example For clarity, we give the transition matrix of a shuffle process, for $n = 3$. We will use $p_k = \mathbb{P}(\pi_1 \to \pi_k) = \mathbb{P}(\text{action } \pi_k)$.

$$\begin{array}{c} \\ (123) \\ (132) \\ (213) \\ (231) \\ (312) \\ (321) \end{array} \left[\begin{array}{cccccc} (123) & (132) & (213) & (231) & (312) & (321) \\ p_1 & p_2 & p_3 & p_4 & p_5 & p_6 \\ p_2 & p_1 & p_5 & p_6 & p_3 & p_4 \\ p_3 & p_4 & p_1 & p_2 & p_6 & p_5 \\ p_5 & p_6 & p_2 & p_1 & p_4 & p_3 \\ p_4 & p_3 & p_6 & p_5 & p_1 & p_2 \\ p_6 & p_5 & p_4 & p_3 & p_2 & p_1 \end{array} \right]$$

Definition Let ϑ be a permutation of length n. This permutation defines a partitioning of the state space in the following manner: π_i and π_j are in the same partition iff there exists an ℓ such that $\pi_i = \vartheta^\ell \pi_j$. We denote the partitions by $P_1, P_2, \ldots, P_m$.

We now consider the shuffling process restricted to a partitioning of the state space, and show that the process of moving from one partition to another is a Markov chain. Our next lemma is a generalization of Theorem 5 of Doner and Uppuluri [2].

Lemma The shuffling process restricted to the partitions is a Markov chain.

Proof We need to show that the transition probabilities only depend on the partitions, from which and to which the transition takes place.

Consider the probability for the transition between partitions P_i and P_j. Let σ be any permutation in P_i and τ any permutation in P_j. Note that $P_i = \{\sigma, \vartheta\sigma, \vartheta^2\sigma, \ldots, \vartheta^{k-1}\sigma\}$, where k is the smallest natural number such that $\vartheta^{k+1} = \vartheta$ (i.e. k is the *order* if ϑ). Similarly, $P_j = \{\tau, \vartheta\tau, \vartheta^2\tau, \ldots, \vartheta^{k-1}\tau\}$. Note that $\mathbb{P}(\sigma \to P_j) = \sum_{\ell=0}^{k-1} \mathbb{P}(\sigma \to \vartheta^\ell \tau) = \sum_{\ell=0}^{k-1} \mathbb{P}(\text{action } (\sigma^{-1}\vartheta^\ell \tau))$. For the other permutations in P_i, we have

$$
\begin{aligned}
\mathbb{P}(\vartheta^r \sigma \to P_j) &= \sum_{\ell=0}^{k-1} \mathbb{P}(\vartheta^r \sigma \to \vartheta^\ell \tau) && (1)\\
&= \sum_{\ell=0}^{k-1} \mathbb{P}(\text{action } (\sigma^{-1}\vartheta^{\ell-r}\tau)) && (2)\\
&= \sum_{\ell'=0}^{k-1} \mathbb{P}(\text{action } (\sigma^{-1}\vartheta^{\ell'}\tau)) && (3)\\
&= \mathbb{P}(\sigma \to P_j). && (4)
\end{aligned}
$$

Therefore, the transition probabilities indeed only depend on the relevant partitions. ■

Note that there are no assumptions needed on ϑ, contrary to the assumptions stated by Doner and Uppuluri [2] in their Theorem 5.

Lemma (Doner and Uppuluri [2]) The eigenvalues of the transition matrix of the new Markov chain form a subset of the eigenvalues of the shuffling process.

We will now show that the transition matrix of the new process is centrosymmetric, when the partitions are ordered in a suitable way. For this, the next lemma is key. We will use the following notation.

Let $R = \{\rho_1, \rho_2, \ldots, \rho_r\}$, where $\rho_i \in \mathcal{S}_n$ for all i. We define

$$\sigma R := \{\sigma\rho_1, \sigma\rho_2, \ldots, \sigma\rho_r\} \tag{5}$$

($R\sigma$ is similarly defined).

Denote by $\Theta := \{\vartheta, \vartheta^2, \ldots, \vartheta^k\}$ (where k is the order of ϑ) the cyclic subgroup with generator ϑ. We recall the definition of the *normalizer* of Θ in $\mathcal{S}_n$, which we will denote by $N_{\mathcal{S}_n}(\Theta)$:

$$N_{\mathcal{S}_n}(\Theta) \quad \stackrel{\text{def}}{=} \quad \{\nu \in \mathcal{S}_n : \nu^{-1}\Theta\nu = \Theta\} \tag{6}$$

Lemma $N_{\mathcal{S}_n}(\Theta) \supsetneq \Theta$ for $n \geq 4$.

Proof We will construct a $\nu \notin \Theta$, such that $\nu^{-1}\Theta\nu = \Theta$. Consider the disjoint cycle decomposition of ϑ. We discern three cases:

(i) Suppose there is a cycle of length $\ell \geq 3$, say $(c_1\ c_2\ \ldots\ c_\ell)$. We construct ν to have the cycles $(c_1\ c_\ell)$, $(c_2\ c_{\ell-1})$, $\ldots$, $(c_{\ell/2}\ c_{\ell/2+1})$ for ℓ even, or $(c_1\ c_\ell)$, $(c_2\ c_{\ell-1})$, $\ldots$, $(c_{(\ell+1)/2-1}\ c_{(\ell+1)/2+1})$, $(c_{(\ell+1)/2})$ for ℓ odd. In other words, we set the c_1th element of ν to c_ℓ and the c_ℓth to c_1, and we set the c_2th element of ν to $c_{\ell-1}$ and the $c_{\ell-1}$th to c_2, etcetera.

For the other cycles $(c'_1\ c'_2\ \ldots\ c'_{\ell'})$, we set the c'_ith element of ν to c'_j, where $j \equiv (k-1)(i-1)+1 \pmod{\ell'}$. Note that this indeed fully prescribes all the elements of ν.

We now have to prove that (1) $\nu \notin \Theta$ and (2) $\nu^{-1}\Theta\nu = \Theta$.

(1) Consider an element in Θ, say ϑ^r. Note the cycle involving c_1 in ϑ^r is $(c_1\ c_{j_1}\ c_{j_2}\ \ldots\ c_{j_\ell})$, where $j_i \equiv ir+1 \pmod{\ell}$. The cylce involving c_1 in ν is $(c_1\ c_\ell)$. For ν and ϑ^r to be equal, these have to be equal, which implies $\ell \equiv r+1 \pmod{\ell}$ and $1 \equiv 2r+1 \pmod{\ell}$. This implies $r \equiv 1 \pmod{\ell}$, but we know for these r, ϑ^r contains the cycle $(c_1\ c_2\ \ldots\ c_\ell)$, which is not equal to $(c_1\ c_\ell)$, since we assumed $\ell \geq 3$.

(2) We'll prove the equivalent statement $\nu\Theta\nu^{-1} = \Theta$. Note that $\nu\vartheta\nu^{-1}$ contains the cycle $(c_\ell\, c_{\ell-1}\ \ldots c_1)$. This, of course, is a cycle of $\vartheta^{\ell-1}$, but also of ϑ^{k-1}, since ℓ is a divisor of k. And by construction, the other cycles of $\nu\vartheta\nu^{-1}$ are also cycle of ϑ^{k-1}. Thus

$$\nu\vartheta\nu^{-1} = \vartheta^{k-1}. \tag{7}$$

This implies $\nu\Theta\nu^{-1} = \Theta$, since

$$\begin{aligned} \nu\vartheta^r\nu^{-1} &= \nu(\nu^{-1}\vartheta^{k-1}\nu)^r\nu^{-1} &\quad (8)\\ &= \vartheta^{r(k-1)} &\quad (9) \end{aligned}$$

and the fact that k and $k-1$ are coprime.

(ii) Suppose there exist at least two cycles of length 2. Let ν be the permutation which only consists of one of these transpositions. Both $\nu \notin \Theta$ and $\nu^{-1}\Theta\nu = \Theta$ are immediately clear.

(iii) Suppose there exist at least two cycles of length 1. Let ν be the permutation which only consists of one transposition of two elements that form cycles of length 1.

Note that for any permutation of length $n \geq 4$, indeed at least one of the above cases applies. ■

Corollary The transition matrix on the partitions is centrosymmetric, when the partitions are ordered in a suitable way.

Proof Let $\bar{M}$ be the transition matrix on the partitions, i.e. $\bar{M}_{ij} = \mathbb{P}(\pi_i \to P_j)$ for any $\pi_i \in P_i$. It suffices to show that there exists a ν such that $\nu P_i = P_{m+1-i}$ for all $i = 1, 2, \ldots, m$, where m is the number of partitions. This is because for any $\pi_i \in P_i$, we have

$$\begin{aligned} \bar{M}_{m+1-i,m+1-j} &= \mathbb{P}(\nu\pi_i \to P_{m+1-j}) \qquad \text{(since } \nu\pi_i \in P_{m+1-i}\text{)} &\quad (10)\\ &= \mathbb{P}(\nu\pi_i \to \nu P_j) &\quad (11)\\ &= \sum_{\sigma\in P_j} \mathbb{P}(\text{action } (\nu\pi_i)^{-1}\nu\sigma) &\quad (12)\\ &= \sum_{\sigma\in P_j} \mathbb{P}(\text{action } \pi_i^{-1}\nu^{-1}\nu\sigma) &\quad (13) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\sigma\in P_j} \mathbb{P}(\text{action } \pi_i^{-1}\sigma) && (14)\\
&= \sum_{\sigma\in P_j} \mathbb{P}(\pi_i \to \sigma) && (15)\\
&= \mathbb{P}(\pi_i \to P_j) && (16)\\
&= \bar{M}_{ij}, && (17)
\end{aligned}
$$

i.e. $\bar{M}$ is centrosymmetric.

Let π_i be any element of P_i. Consider a $\nu \in N_{\mathcal{S}_n}(\Theta)\backslash\Theta$. Note that (by definition of $N_{\mathcal{S}_n}(\Theta)$), we have $\nu\Theta = \Theta\nu$. Therefore

$$\nu P_i = \nu\Theta\pi_i = \Theta\nu\pi_i = \Theta\pi_j, \tag{18}$$

for some $\pi_j \notin P_i$, since $\nu \notin \Theta$.

The above gives a recipe for creating a centrosymmetric matrix: Start with any partition P_i (permutation π_i). Its "antipode" is νP_i ($\nu\pi_i$) (this is trivially unique for each partition). Fix the position (number) of these partitions, and take a new partition, that is not fixed yet. Repeat this process, until the order of the partitions is found. ■

Summarizing, we have devised a method to find a Markov chain with a smaller state space (of size $n!/k$, where k is the order of ϑ), which has eigenvalues which form a subset of the eigenvalues of the shuffle process. Furthermore, the transition matrix of this new Markov chain can be made centrosymmetric, and therefore, using the method from Collar [1], we can consider two matrices of size $(n!/(2k) \times n!/(2k))$, when computing the eigenvalues.

The question is which ϑ's to use. It turns out that we don't have to consider all $\vartheta \in \mathcal{S}_n$:

Lemma ϑ and $\sigma\vartheta\sigma^{-1}$ give rise to the same transition matrix (except for reordering), for all σ.

Proof Note that the partitions induced by $\sigma\vartheta\sigma^{-1}$, are exactly σP_1, σP_2, $\ldots$, σP_m, where $P_1, P_2, \ldots, P_m$ are the partitions induced by ϑ. Now it is easy to see that the shuffling actions that transform a permutation in P_i to the permutations in P_j, will also transform a permutation in σP_i to the permutations in σP_j. The transition matrices will therefore be same, when suitably ordered. ■

Recall the wellknown fact that two permutations are conjugates, iff the sets of cycle lengths of both permutations are equal. This implies that we only need to consider all possible ϑ's with different cycle lengths to create the smaller matrices.

An interesting open question is whether *all* eigenvalues of the shuffle matrix can be obtained by just looking at the eigenvalues of the transition matrix of the "partition processes" for all ϑ not equal to the identity.

3 Summary

We proposed a method for computing eigenvalues of a shuffle matrix in an easier way, by looking at a Markov chain on a collapsed state space. In our exposition of the method, we show that $N_{\mathcal{S}_n}(\Theta) \subsetneq \Theta$ for $n \geq 4$, a fact that the authors have not encountered in the literature.

4 Acknowledgements

We would like to thank Robert van der Waall for his indispensable help in proving $N_{\mathcal{S}_n}(\Theta) \subsetneq \Theta$ for $n \geq 4$. We are indebted to Professor Henk Tijms for suggesting "shuffling cards" as the subject of our thesis.

References

[1] A. R. Collar. On centrosymmetric and centroskew matrices. *The Quarterly Journal of Mechanics and Applied Mathematics*, 15:265–281, 1962.

[2] J. R. Doner and V. R. R. Uppuluri. A Markov chain structure for riff shuffling. *SIAM Journal on Applied Mathematics*, 18:191–209, 1970.