Contemporary Mathematics

The Covariance Structure of Random Permutation Matrices

Marlos A. G. Viana

Abstract. Random permutation matrices (U) arise naturally in the stochastic representation of vectors of order statistics, induced order statistics and associated ranks. When the probability law of U is uniform, the covariance structure among the entries of U is derived explicitely, and a constructive derivation for the covariance in the general case (U^k) is described and related to the cyclic structure of the symmetric group. It is shown that the covariance structure of vectors resulting from the multiplicative action UX of U on a given vector X requires averaging the symmetric conjugates $UXX'U'$ of XX' over the group of permutation matrices. The mean conjugate is a projection operator which leads to a trace decomposition with the usual ANOVA interpretation. Numerical examples of these decompositions are discussed in the context of the analysis of circularly symmetric data.

1. Introduction

Permutation matrices are square matrices characterized by having all entries either 0 or 1 and exactly one 1 in each row and column. The collection (G) of all permutation matrices (g) of order n is in one-to-one correspondence (ρ) with the collection (S_n) of all permutations (π) on a set (L) of n objects, like $\{1, 2, \dots, n\}$. In fact, the permutation matrix associated with a permutation π is the matrix representing the linear transformation that maps the canonical basis $\{e_1, \dots, e_n\}$ of R^n into the basis $\{e_{\pi(1)}, \dots e_{\pi(n)}\}$. The collection S_n together with the operation of function composition defines the symmetric group on n objects. The correspondence ρ is defined in S_n with values in the space (GL_n) of invertible square matrices of order n over the field of complex numbers, with the homomorphic property $\rho(\pi_1\pi_2) = \rho(\pi_1)\rho(\pi_2)$ for all π_1, π_2 in S_n. This characterizes ρ as a n-dimensional linear representation of S_n (the permutation representation). If a permutation π is selected according to a probability law in S_n, then $\rho(\pi)$ is a random permutation matrix (U) in G. In the present paper we concentrate on random permutation matrices in G, with the perspective that those are simply one of the many (random) linear representations for random events defined in the symmetric group S_n. Unless for clarification, we will write G to indicate either S_n or M_n.

1991 *Mathematics Subject Classification.* 15A52, 20B30.
Key words and phrases. random matrices, symmetric group, moments.

Random permutation matrices arise naturally, for example, in the linear representation of order statistics. This is illustrated in the introductory examples below, where the determination of the uniform mean $E(U)$ and the uniform mean conjugate $E(UHU')$ of a given matrix H are required. The uniform mean is simply the average of all elements in G, whereas the mean conjugate of H is defined as the average of the matrix conjugates gHg' of H over G. More generally, as the examples will show, we need to describe the corresponding uniform higher-order means $E(U^k)$ and mean conjugates $E(U^kHU^{k'})$. This is equivalent to deriving the covariance structure among the entries of U^k.

The covariance structure of U, a uniform random permutation matrix, is studied in detail in Section 2. Proposition 2.2 in Section 2 states that the uniform mean conjugate of H decomposes as $v_0ee'/n + v_1(I - ee'/n)$, where $v_0 = e'He/n$ and $v_0 + (n-1)v_1 = \text{tr } H$. This result, when applied to the linear representation of order statistics, for example, leads to the covariance structure of ranks of permutation symmetric observations, or to the mean and variance of the usual Wilcoxson rank-sum statistics. Proposition 2.2 is a direct consequence of the (n-1)-dimensional irreducible linear representation of G, which leads to a decomposition of tr $E(UHU')$ into the weighted average of corresponding orthogonal projections with weights proportional to the dimensions of the irreducible subspaces, and is carried out by the intertwining action of $E(UHU')$. The ANOVA interpretation of such decomposition is discussed in Section 3, where an application to circularly symmetric data is outlined. Finally, in Section 4 we derive the covariance structure of moments of random permutations. The main result (Theorem 4.1) shows that, when the law of U is uniform in G, the trace of U^k is in average equal to the number of positive integer divisors of k not greater than n.

Example 1.1. Representation of order statistics. Let X_i indicate the rank of the i-th component Y_i of a random vector Y, defined on a linear subspace V of $\mathbb{R}^n$, when these components are ordered from the smallest to the largest value, and assume for the moment that either the components of Y are distinct or that there is a fixed rule that assigns distinct ranks to tied observations. Let $\pi \in G$ indicate the permutation taking X_i to i, that is, $X_{\pi(i)} = i$, $i = 1, \dots, n$. Given the standard basis $\{e_1, \dots, e_n\}$ of V, define the $n \times n$ matrix $\rho(\pi)$ by $e_{\pi(j)} = \sum_i \rho(\pi)_{ij} e_i, \quad j = 1, \dots, n$. Equivalently, $\rho(\pi)_{ij} = 1$ if and only if π takes j into i, that is $\pi(j) = i$. The homomorphic map $\pi \to \rho(\pi)$ is the *permutation representation* introduced above. By virtue of the randomness in the ranks of Y, the matrix $U = \rho(\pi)$ is a random permutation matrix, and the multiplicative action $U'Y$ of U' on Y linearly represents the ordered version of Y. For example, if the vector of ranks of an observed vector y is $x' = (3, 4, 2, 1)$, then π is the permutation taking x' into $r' = (1, 2, 3, 4)$. The cycle notation for π is (1423), so that $\pi : 1 \to 4$, $4 \to 2$, $2 \to 3$ and $3 \to 1$. The corresponding linear representation is

$$u = \rho(\pi) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

If we indicate by $\mathcal{Y}$ the ordered version of Y, then the random permutation matrix U ($= \rho(\pi)$) representing the permutation $\pi : X \to r$ is such that $X = Ur$ and $\mathcal{Y} = U'Y$. If (Y,Z) are concomitant random vectors of corresponding dimensions, then the action U'Z of U' on Z generates the vector of concomitants of order (or induced order) statistics. If the distribution of Y is permutation symmetric (that is Y and gY are equally distributed for all $g \in G$), then the probability law of U is uniform in G. The covariance structure of ranks, order statistics and concomitants may then be expressed as multiplicative actions of uniformly distributed random permutations. In particular, to obtain the covariance structure of ranks $X = Ur$, we need to evaluate the uniform expected value $\sum_g gr/n!$ of $U'r$, indicated by $E(Ur)$, and the uniform expected value $\sum_g grr'g'/n!$ of $U'rr'U$, denoted by $E(Urr'U')$. The characterization and interpretation of uniform averages over G, such as those in $E(U)$ and $E(Urr'U')$ will be obtained in Section 2.

Other formulations leading to random linear representations can be obtained in terms of random walks or random graphs, or more generally as iterated random functions (e.g., [**DF99**]).

2. Means and mean conjugates

Let U indicate a random permutation matrix distributed according to $w \equiv \{w(g); g \in G\}$. We refer to

$$E_w(UHU') = \sum_{g \in G} gHg^{'} w(g)$$

as the mean conjugate of H when U is distributed according to w on G. The mean conjugate is a particular invariant mean $\sum_g c(g)/|G|$ on G, where c is a function defined in G. [e.g., [**Nai82**, p.60]]. When U is uniformly distributed in G we write $E(UHU')$ and refer to the uniform mean conjugate of H. When an operator M commutes with every element of G we say that M *has the symmetry of* G.

PROPOSITION 2.1. *$E_w(UHU')$ has the symmetry of G if and only $E_w(UHU')$ is the uniform mean conjugate of H.*

PROOF. If $E_w(UHU')$ is the uniform mean conjugate of H, the transitivity of G shows that $tE(UHU')t' = E(UHU')$ for all $t \in G$. Conversely, if U is assigned the value g with probability w(g), and $E_w(UHU')$ commutes with every element t of G, we have $E_w(UHU') = t\sum_g w(g)gHg't' = \sum_g w(g)tgH(tg)' = \sum_v w(t'v)vHv'$, where the last equality uses the transitivity of G. The same fact shows that $\sum_t w(t'v) = 1$, for all t, so that

$$E_w(UHU') = \frac{1}{n!}\sum_t \sum_v w(t'v)vHv' = \sum_v \frac{1}{n!}\sum_t w(t'v)vHv' = \sum_v \frac{1}{n!}vHv'.$$

□

The following result gives the explicit evaluation of $E(UHU')$ (see also [**Dan62**] for an heuristic proof).

PROPOSITION 2.2. *(a) The uniform mean E(U) of U in G is ee'/n and (b) the uniform mean conjugate $E(UHU')$ of H on G decomposes as $v_0\frac{ee'}{n} + v_1(I - \frac{ee'}{n})$, where $v_0 = e'He/n$ and $v_0 + (n-1)v_1 = tr\ H$.*

PROOF. Let V be the one-dimensional subspace of $\mathbb{R}^n$ spanned by $e' = (1, ..., 1)$, W its orthogonal complement in $\mathbb{R}^n$, $M = E(U) - \frac{ee'}{n}$ and $\mathcal{M} = E(UHU') - v_0\frac{ee'}{n} - v_1(I - \frac{ee'}{n})$, with v_0 and v_1 as above. Allowing for complex solutions, if all eigenvalues of M or $\mathcal{M}$ are zero then the proposition holds. Otherwise, let λ be a non-zero eigenvalue of M, ℓ a non-zero eigenvalue of $\mathcal{M}$, $K = kern\ \{M - \lambda I\}$ and

$\mathcal{K} = kern\ \{\mathcal{M} - \ell I\}$. We will show that both K and $\mathcal{K}$ are subspaces of W stable under G, and use the fact[e.g., [**Ser77**, p.17], [**Dia88**, p.6] and Appendix A] that V and W are irreducible subspaces of the linear representation ρ. Consequently, both K and $\mathcal{K}$ must reduce to null subspaces, and this will complete the proof.

(1) $K \subset W$. First note that $e'M = e'(\frac{1}{n!}\sum g - \frac{ee'}{n}) = 0$, so that for $v \in K$ we have $(M - \lambda I)v = 0 = e'M - \lambda e'v = \lambda e'v$, and because $\lambda \neq 0$ we obtain $e'v = 0$, or $v \in W$.
(2) M has the symmetry of G. In fact, from the transitivity of G, $tMt' = \frac{1}{n!}\sum tgt' - \frac{tee't'}{n} = \frac{1}{n!}\sum_g g - \frac{ee'}{n} = M$, $\forall t \in G$.
(3) K is stable under G. Take $v \in K$ and $t \in G$, then (by 2) $Mtv = tMv = t\lambda v = \lambda tv$, that is $tv \in K$, for all $t \in G$. This completes the proof of part (a).
(4) $\mathcal{K} \subset W$. We use part (a) to show, similarly to 1, that $e'\mathcal{M} = 0$. In fact, $e'\mathcal{M} = \frac{1}{n!}\sum_g e'gHg' - v_0\frac{e'ee'}{n} - v_1(e' - \frac{e'ee'}{n}) = e'H\frac{1}{n!}\sum_g g - v_0 e' = e'H\frac{1}{n}ee' - v_0e' = (\frac{1}{n}e'He - v_0)e' = 0$, by the definition of v_0. Now for $v \in \mathcal{K}$ we have $(\mathcal{M} - \ell I)v = 0 = e'\mathcal{M} - \ell e'v = \ell e'v$, and because $\ell \neq 0$ we obtain $e'v = 0$, or $v \in W$.
(5) $\mathcal{M}$ has the symmetry of G. Again, from the transitivity of G, $t\mathcal{M}t' = \sum_g tgHg't' - v_0\frac{tee't'}{n} - v_1(tt' - \frac{tee't'}{n}) = t\mathcal{M}t'$, $\forall t \in G$.
(6) $\mathcal{K}$ is stable under G. Take $v \in \mathcal{K}$ and $t \in G$, then (by 5) $\mathcal{M}tv = t\mathcal{M}v = t\ell v = \ell tv$, that is $tv \in \mathcal{K}$, for all $t \in G$. This completes the proof of part (b).

□

Equivalently, the decomposition of $E(UHU')$ may be expressed as $a_0ee' + a_1I$, for some scalars a_0 and a_1 determined by $n(a_0 + a_1) = \text{tr } H$ and $n(n-1)a_0 = e'He - \text{tr } H$.

EXAMPLE 2.1. Let U indicate the order statistics representation introduced in Section 1 associated with a random vector Y, and let $r' = (1, 2, \ldots, n)$. The vector $U'r$ indicates the vector of ranks associated with Y, when the components of Y are all distinct. Average ranks with tied observations are discussed in Example 2.3. If the distribution of Y is permutation symmetric then U acts uniformly on r and $U \sim U'$ in distribution. The vector of means, $E(Ur)$, from Proposition 2.2, is

$$E(Ur) = E(U)r = \frac{\mathbf{ee}'}{n}r = \frac{n(n+1)}{2n}\mathbf{e} = \frac{n+1}{2}\mathbf{e},$$

whereas the covariance of Ur is given by $E(Urr'U') - E(Ur)E(Ur)'$. To evaluate the mean conjugate of rr' in G, following Proposition 2.2, we note that the trace of rr' is $n(n+1)(2n+1)/6$, the sum of the squares of the first n positive integers whereas the sum $e'rr'e$ of all entries of rr' is $(\mathbf{e}'r)^2 = [n(n+1)/2]^2$. Therefore,

$$n(a_0 + a_1) = \frac{n(n+1)(2n+1)}{6},$$
$$n(n-1)a_0 = [\frac{n^2(n+1)^2}{4} - \frac{n(n+1)(2n+1)}{6}] = \frac{1}{12}n(n+1)(n-1)(3n+2),$$

from which we obtain

$$\text{Cov}(Ur) = [a_0 - \frac{(n+1)^2}{4}]\mathbf{ee}' + a_1\mathbf{I} = \frac{-(n+1)}{12}\mathbf{ee}' + \frac{n(n+1)}{12}\mathbf{I}. \tag{2.1}$$

The result shows that the common variance among the ranks is $(n+1)(n-1)/12$, and the common covariance between any two ranks is $-(n+1)/12$. The resulting common correlation is -1/(n-1). The mean and variance of the usual Wilcoxson rank-sum statistics W follows from fixing (say, the first) m components of R. Writing $f' = (1, ..., 1, 0, ..., 0)$ with $1 \leq m < n$ components equal to 1, then $W = f'R$ and $EW = f'E(Ur) = m(n+1)/2$, whereas $var(W) = f'\text{Cov}(Ur)f = m(n-m)(n+1)/12$. Similar arguments can be applied to rank correlations (e.g., [**BE73**]).

Example 2.2. Covariance structure of order statistics. Consider again the linear representation $U'Y$ of the order statistics $\mathcal{Y}$. The covariance structure of $\mathcal{Y}$ requires computing $E(\mathcal{Y})$ and $E(\mathcal{Y}\mathcal{Y}')$. Suppose we want to evaluate $E(\mathcal{Y})$ in terms of $U'Y$. Let $\Gamma_1 = \{y \in \mathbb{R}^n : y_1 \leq \ldots \leq y_n\}$, denote by Γ_g the set image $g\Gamma_1$ of Γ_1 under the permutation matrix g, and set $w(g) = P(Y \in \Gamma_g)$. Let also $E_g Y$ indicate the expectation of Y conditional to Γ_g, that is, $E_g Y = \int y I_g(y) F(dy)/w(g)$, where I_g is the indicator function of Γ_g and F is the probability law of Y. The expectation of $\mathcal{Y} = U'Y$ is then

$$E(U'Y) = \sum_{g \in G} w(g) g' E_g Y = \sum_{g \in G} \int g' y I_g(y) F(dy) = \sum_{g \in G} \int x I_g(gx) F(gdx).$$

which reduces to

$$E(U'Y) = \int x I_1(x) F^*(dx), \tag{2.2}$$

where $F^*(dx) = \sum_{g \in G} F(gdx)$. If, in addition, F is a permutation symmetric measure (as in the usual iid case), $E(U'Y)$ reduces to the usual form $\int x I_1(x) n! F(dx)$. Note that equation (2.2) is a template to natural extensions. For example, consider the case of conditioning Y on a concomitant (eventually random) vector X. Given X=x, we observe Y_x according to the probability law of $Y|x$ and represent its ordered version by $U'_{Y_x} Y_x$, as before. By definition, we have $E(U'Y \mid x) = E(U'_{Y_x} Y_x)$, $w_x(g) = P(Y \in \Gamma_g | x)$, so that

$$E(U'Y) = E_x E(U'Y \mid x) = E_x \sum_{g \in G} w_x(g) g' E_g(Y_x).$$

Conditioning is attractive when the probability law of Y_x is symmetrically generated, e.g., $Y_x = Mx + V$ where M and V have the symmetry of G. The covariance structure of ordered observations from symmetrically dependent observations is described in [**LV99**], [**Via98**], [**VO97**], [**OV95**].

Example 2.3. The representation of ordered tied observations. Given $y \in \mathbb{R}^n$, define the equivalence relation on $\{1, \ldots, n\}$: $i \equiv j$ if and only if $y_i = y_j$, and let $C_1, \ldots, C_m$ be the resulting equivalent classes, the C_i class with k_i elements and block-rank i. The k_i rank-equivalent elements in C_i are assigned to labels $c_{i1}, \ldots, c_{ik_i}$, for example, monotonically increasing with j. Index the canonical basis accordingly by e_{ij}, $i = 1, \ldots, m$ and $j = 1, \ldots, k_i$. Let π_i be a permutation in $S_{k_i} \equiv S_i$, $i = 1, \ldots, m$ and define the action $e_{ij} \to e_{\pi_i(c_{ij})}$ of the product group $\Pi_i S_i$ on the standard basis. Let $\rho(\pi_1, \ldots, \pi_m)$ be the corresponding permutation representation. Then, for any choice of $\pi_1, \ldots, \pi_m$, $\rho(\pi_1, \ldots, \pi_m)$ is an order representation, that is, $\rho(\pi_1, \ldots, \pi_m)Y = \mathcal{Y}$, whereas the resulting vector R of average ranks is obtained from the multiplicative action of the average inverse order representations on $r' = (1, 2, \ldots, n)$, that is,

$$R = \frac{1}{\Pi_i k_i!} \sum_{S_1 \times \cdots \times S_m} \rho(\pi_1, \ldots, \pi_m)' r.$$

To illustrate, let $y' = (1, 0, 0, 1, -1)$. Then $C_1 = \{5\}$, $C_2 = \{2, 3\}$ and $C_3 = \{1, 4\}$, so that $c_{11} = 5$, $c_{21} = 2$, $c_{22} = 3$, $c_{31} = 1$ and $c_{32} = 4$. The resulting actions

$$e_{11} \to e_{\pi_1(c_{11})}, e_{21} \to e_{\pi_2(c_{21})}, e_{22} \to e_{\pi_2(c_{22})}, e_{31} \to e_{\pi_3(c_{31})}, e_{32} \to e_{\pi_3(c_{32})},$$

of $S_1 \times S_2 \times S_2$ on the standard basis leads to the order representations. Corresponding to the identity, for example, a linear order representation U is

$$U = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

with the effect that $UY = \mathcal{Y}$. There are $1!2!2! = 4$ of those and their sum is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 4 \\ 0 & 2 & 2 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 & 0 \end{bmatrix}.$$

Taking the transpose, it follows that,

$$\frac{1}{4} \begin{bmatrix} 0 & 0 & 0 & 2 & 2 \\ 0 & 2 & 2 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 4 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 4.5 \\ 2.5 \\ 2.5 \\ 4.5 \\ 1 \end{bmatrix}.$$

EXAMPLE 2.4. Familial/Bilateral Index. The argument is proposed in [**PGL$^+$88**]: If familial factors do not influence the risk of (certain types of) cancer beyond their effects on sex, age at treatment and dose, then the residuals (on a proportional hazards model, say) should be distributed independently of family membership. The proposed statistics is the sum of the products of the residuals $\{x_i, y_i\}$ for pairmates $i = 1, \ldots, n$. Large values of $T = \sum_i^n x_i y_i = x'y$ are indicative of familial effects. To assess the expected variance of T, given the data x,y, under the assumption of random pairing of siblings, define the random outcome $T_u = x'Uy$, where U is a random permutation matrix, uniformly distributed in G. Then, $var(T_u) = x'E(Uyy'U')x$, and direct application of Proposition 2.2 shows that $E(T_u) = n\bar{x}\bar{y}$, and that $var(T_u) = (n-1)S_x^2 S_y^2$, where S_x^2 and S_y^2 are the corresponding usual estimates of the lateral variability in the data.

3. Trace decompositions

Proposition 2.2 shows that the irreducible $V \oplus W$ decomposition of G leads to a decomposition of tr $E(UHU')$ into the weighted average of corresponding orthogonal projections with weights proportional to dim V and dim W, respectively.

TABLE 1. Trace decomposition for the (n-1) irreducible representation of G.

effect	components	df	mss	$\mathcal{F}$
W	$(n-1)v_1$	n-1	v_1	v_1/v_0
V	v_0	1	v_0	
total	tr H	n		

This is carried out by the intertwining action of $E(UHU')$. Table 1 illustrates the decomposition and its ANOVA interpretation ($v_0 \neq 0$).

EXAMPLE 3.1. The following cases will justify the interpretation:

(1) Consider the case H=mI, $m \neq 0$. Then $v_0 = v_1 = m$ and $\mathcal{F} = 1$;

(2) Let $H \in G$, uniformly. Then $v_0 = e'He/n = 1$ and $\mathcal{F} = v_1 = (\text{tr }(H) - 1)/(n-1)$. One shows that $\mathcal{F}$ has mean zero and standard deviation $1/(n-1)$ under uniform sampling in G. Moreover, $\mathcal{F} = 1$ if and only if H=I and the whole n-dimensional $\mathbb{R}^n$ is invariant, whereas $\mathcal{F} = -1/(n-1)$ if and only if H is cyclic and the only largest invariant subspace is the 1-dimensional subspace generated by $e' = (1, ...1)$ (the image of ee'), so that larger values of $\mathcal{F}$ reflect a larger invariant subspace;

(3) Let H be doubly stochastic with non-negative entries. It can be written as a convex combinations of permutation matrices, say, $H = w_1 g_1 + \ldots + w_r g_r$. Direct computation shows that

$$\mathcal{F} = \frac{\text{tr } H - 1}{n-1}.$$

Again, because H is non-negative and doubly stochastic, large values of $\mathcal{F}$ imply that the off-diagonal entries must be very small and hence H must be concentrated on its diagonal. In fact $\mathcal{F} = 1$ if and only if H=I.

(4) Let $y \in \mathbb{R}^n$ and $H = yy'$. Then $(n-1)v_1 = ||y - \bar{y}e||^2$ and $v_0 + (n-1)v_1 = ||y||^2$, so that

$$\mathcal{F}^{-1/2} = \sqrt{n}\frac{|\bar{y}|}{\sqrt{\sum_i (y_i - \bar{y})^2/(n-1)}}.$$

(5) Let $H = \sum_i (x_i - \bar{x})(y_i - \bar{y})'$. Then,

$$\mathcal{F} = \frac{1-r}{1+(n-1)r},$$

where r is the usual sample correlation coefficient under the permutation symmetric assumption for the underlying covariance matrix, that is, under the assumption that the covariance matrix has the symmetry of G. More precisely,

$$r = \frac{e'He - \text{tr } H}{(n-1)\text{tr } H}.$$

(6) Let $y \in \{0,1\}^n$, $H = yy'$ and $\bar{y}$ the proportion of ones in y. Then one shows that

$$\mathcal{F} = (n-1)\frac{\bar{y}}{1-\bar{y}},$$

that is, $\mathcal{F}/(n-1)$ is the observed odds on "1"-components in y.

Example 3.2. Cyclic decomposition. Let g in G be an element of order n, and let C_n indicate the corresponding cyclic group generate by g. In direct analogy to Proposition 2.2, we obtain that the uniform mean conjugate $E(UHU')$ of H on C_n decomposes as $v_0 I + v_1 g + \ldots + v_{n-1} g^{n-1}$, where the weights v_i are given by $v_i = \text{tr } g^{n-i} H/n$. When H is symmetric, we have, in addition, that $v_i = v_{n-i}$. The corresponding trace decomposition ANOVA for n=4 (and similarly for any even n) and $H = yy'$ is illustrated on Table 2. The case n=5 (and similarly for any odd n) is illustrated on Table 3.

Table 2. Analysis of Variance for Cyclic Decomposition, n=4, $H = yy'$.

i	components	df	mss (v_i)
0	tr $H/n = y'y/4$	1	$y'y/4$
1	2tr $gH/n = 2y'gy/4$	2	$y'gy/4$
2	tr $g^2H = 1y'g^2y/4$	1	$y'g^2y/4$
total	$e'He/n = 4\bar{y}^2$	4	

Table 3. Analysis of Variance for Cyclic Decomposition, n=5, $H = yy'$.

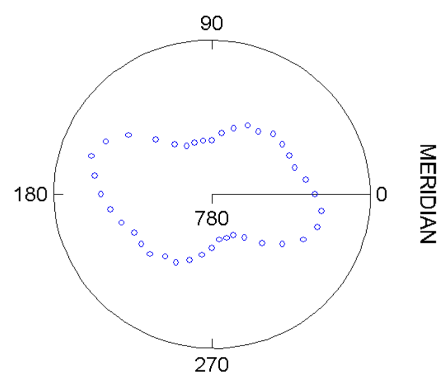
i	components	df	mss (v_i)
0	tr $H/n = y'y/5$	1	$y'y/5$
1	2tr $gH/n = 2y'gy/5$	2	$y'gy/5$
2	2tr $g^2H = 2y'g^2y/5$	2	$y'g^2y/5$
total	$e'He/n = 5\bar{y}^2$	5	

Example 3.3. ANOVA for cyclic decompositions (Keratometry Data). Keratometry is the measurement of corneal curvature of a small area using a sample of four reflected points of light along an annulus 3 to 4 mm. in diameter, centered about the line of sight. For normal cornea this commonly approximates the apex of the cornea [[**VOM93**]]. The fundamental principle of computerized keratometry is similar in that the relative separation of reflected points of light along concentric rings are used to calculate the curvature of the measured surface. Using a pattern of concentric light-reflecting rings and sampling at specific circularly equidistant intervals, a numerical model of the measured surface may be obtained. Sampling takes place at equally-spaced ring-semimeridian intersections [e.g., [**KW89**]]. Figure 1 shows the apex (scaled) distances measured along 360 equally-spaced semimeridians (9-degree separation, n=40), based on the reflected image of a fixed ring. Similar data are generated for the actual surface curvature. The resulting ANOVA is shown in Table 4. The total variability $40 \times \bar{y}^2 = 0.2498454423\ 10^8$ is decomposed into the cyclic components $y'g^iy/40$. More specifically, $40 \times \bar{y}^2 = \sum_{i=0}^{40} v_i \times df_i$. Figure 2 illustrates the size of the semimeridian *effects*, with maximum effect sizes orthogonally located. The analysis also picks up the direction (corresponding to steep and flat curvatures) of the corresponding effects.

TABLE 4. Analysis of Variance for Circular Data with decomposition $40 \times \bar{y} = 0.2498454423\ 10^8 = \sum_{i=0}^{40} y' g^i y$.

meridian (i)	$y'g^iy$	df_i	meridian (i)	$y'g^iy$	df_i
0	2501.44	1	10, 30	2496.61	2
1, 21	2501.19	2	11, 31	2496.63	2
2, 22	2500.56	2	12, 32	2496.72	2
3, 23	2499.70	2	13, 33	2496.93	2
4, 24	2498.80	2	14, 34	2497.30	2
5, 25	2498.00	2	15, 35	2497.88	2
6, 26	2497.39	2	16, 36	2498.66	2
7, 27	2496.99	2	17, 37	2499.55	2
8, 28	2496.76	2	18, 38	2500.39	2
9, 29	2496.65	2	19, 39	2501.00	2
			20	2501.23	1
			total	$0.2498454423\ 10^8$	40

FIGURE 1. Keratometry data with 9-degree meridian separation.



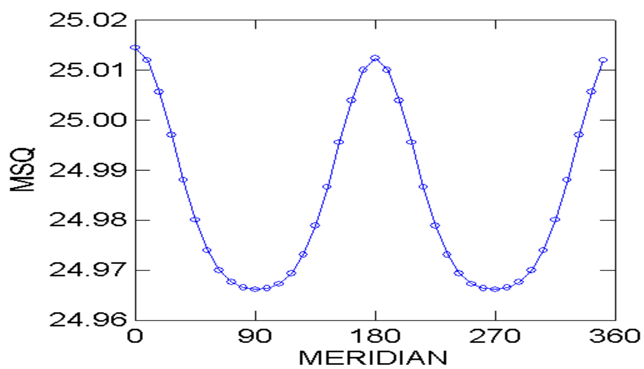
4. The covariance structure and transition probabilities

Recall that given two random matrices U and V, the (ij,st)-entry of $\text{Cov}(UV)$ is given by $\text{Cov}(UV)_{ij;st} = e_i^{'}\text{E}(UV^{'}e_je_t^{'}VU^{'})e_s - e_i^{'}\text{E}(UV)e_je_t^{'}\text{E}(UV)^{'}e_s$. Taking $m = n$ and $\mathbf{V}$ constant and equal to $\mathbf{I}$, the (ij,st) entry of $\text{Cov}(U^k)$ is expressed as

$$\text{Cov}(U^k)_{ij,st} = E(e_i^{'}U^ke_je_s^{'}U^ke_t) - E(e_i^{'}U^ke_j)E(e_s^{'}U^ke_t), \quad k = 1, \ldots . \tag{4.1}$$

The ij-entry $U_{ij,k} = e_i^{'}U^ke_j = e_i^{'}e_{U^k(j)}$ of U^k indicates whether or not U^k moves j into i. That is, observed U=g, the value of $U_{ij,k}$ is 1 if $g^k(j) = i$ and is 0 otherwise. Then, under the uniform law, the average $E(U^k)$ of U^k over G is the matrix in which the ij-component is the probability $P_k(j \to i)$ that U^k moves j into i. Similarly, $U_{ij,k}U_{st,k}$ indicates whether or not U^k moves (j,t) into (i,s). The corresponding

FIGURE 2. Analysis of Variance for Circular Data, $MSQ = y'g^iy/10^6$, for each semimeridian angle 9 i, $i = 1, \ldots, 40$.



probability $\mathcal{P}_k((j,t) \to (i,s))$ is the average $E(U_{ij,k}U_{st,k})$ of $U_{ij,k}U_{st,k}$ over G under the uniformly probability law. It then follows that the (ij,st)-entry (4.1) of Cov(U^k) has the interpretation

$$\text{Cov}(U^k)_{ij,st} = \mathcal{P}_k[(j,t) \to (i,s)] - P_k(j \to i)P_k(t \to s).$$

Denoting the associated random walk on $\{1, \ldots, n\}$ by L (with law P_k) and the random walk on $\{1, \ldots, n\}^2$ by L^2 (with law $\mathcal{P}_k$), we write, shortly, $\text{Cov}(U^k) = \mathcal{P}_k(L^2) - P_k(L)^2$. There is also a random walk in the dual tensor space generated by $e_i e'_{U(j)}$, which is also of interest. The entries of $\mathcal{P}_k(L^2)$ are generally assembled in lexicographic order $11, \ldots, 1n, \ldots, n1, \ldots, nn$: Let $C_{l,k}$ indicate the l-th column of U^k and C_k the $1 \times n$ block matrix $(C_{1,1}, \ldots, C_{n,k})$. Then,

$$\text{Cov}(U^k) = E(C_k C'_k) - E(C_k)E(C_k)', \tag{4.2}$$

with the (l,m) block corresponding to $\text{Cov}(C_{l,k}, C_{m,k})$, $l, m = 1 \ldots, n$, displays the covariance structure of U^k in lexic form.

4.1. The covariance structure of U. The following proposition describes the covariance structure of a random permutation matrix.

PROPOSITION 4.1. *If U is uniformly distributed in G, then,*

$$Cov(U)_{ij;st} = \frac{1}{n^2(n-1)}[1 + n(n\delta_{jt}\delta_{is} - \delta_{jt} - \delta_{is})]. \tag{4.3}$$

PROOF. Taking expression (4.1) with k=1, E($Ue_je'_tU'$) is the uniform mean conjugate of $e_je'_t$, and hence, from Proposition 2.2, it has the form $a_0\mathbf{e}\mathbf{e}' + a_1I$, with

$$n(n-1)a_0 = \text{tr }(\mathbf{e}\mathbf{e}' - I)e_je'_t = 1 - \delta_{jt}, \quad n(a_0 + a_1) = \text{tr } e_je'_t = \delta_{jt}.$$

Substituting the expressions for $E(Ue_je'_tU')$ and $E(U) = \frac{1}{n}\mathbf{e}\mathbf{e}'$ into (4.1), the result follows. □

EXAMPLE 4.1. To illustrate with $n = 3$, the covariance structure of U is determined by the covariance between any two columns i and s;

$$\mathrm{Cov}(C_i, C_s) = \frac{1}{18}\left[\begin{array}{rrr|rrr} 4 & -2 & -2 & -2 & 1 & 1 \\ -2 & 4 & -2 & 1 & -2 & 1 \\ -2 & -2 & 4 & 1 & 1 & -2 \\ \hline -2 & 1 & 1 & 4 & -2 & -2 \\ 1 & -2 & 1 & -2 & 4 & -2 \\ 1 & 1 & -2 & -2 & -2 & 4 \end{array}\right].$$

In general, the covariance matrix between any two columns i and s of U with uniform distribution in G has the form

$$\mathrm{Cov}(C_i, C_s) = \begin{cases} \frac{1}{n^2(n-1)}\mathbf{e}\mathbf{e}' + \frac{-1}{n(n-1)}\mathbf{I}, & \text{when } i \neq s, \\ \frac{-1}{n^2}\mathbf{e}\mathbf{e}' + \frac{1}{n}\mathbf{I}, & \text{when } i = s. \end{cases}$$

4.2. The covariance structure of U^k. Here it will be convenient to assemble the entries of $\mathrm{Cov}(U^k)$ using the matrix direct-product $U \otimes V$ notation. We also write $\otimes^2 V$ to indicate $V \otimes V$. First we note that both $E(\otimes^2 U^k)$ and $\otimes^2 E(U^k)$ are block matrices, with respective blocks $E(U_{ij,k}U^k)$ and $E(U_{ij,k})E(U^k)$ indexed by ij. Within each block the entries are indexed by s (rows) and t (columns). Clearly, $\mathrm{Cov}(U^k) = E(\otimes^2 U^k) - \otimes^2 E(U^k)$. For example, the sums of the rows in lexicographic notation, as in (4.2), correspond to block sums in $\otimes$-notation. To illustrate, with n=2, we have (omitting the k-index),

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}, \quad \otimes^2 U = \left[\begin{array}{cc|cc} U_{11}U_{11} & U_{11}U_{12} & U_{12}U_{11} & U_{12}U_{12} \\ U_{11}U_{21} & U_{11}U_{22} & U_{12}U_{21} & U_{12}U_{22} \\ \hline U_{21}U_{11} & U_{21}U_{12} & U_{22}U_{11} & U_{22}U_{12} \\ U_{21}U_{21} & U_{21}U_{22} & U_{22}U_{21} & U_{22}U_{22} \end{array}\right].$$

The associated random walk L^2 visits four points labeled by (1,1),(1,2),(2,1) and (2,2). The transition from (j,t) to (i,s) occurs with probability $E(U_{ij}U_{st})$. The diagonal of the matrix $E(\otimes^2 U)$ contains the *resting* probabilities $[(i,j) \to (i,j)]$ and the average trace of $E(\otimes^2 U)$, for example, is the walk's resting probability when the past state distribution is uniform. Similarly, $E(U)$ describes the marginal transition probabilities associated with the marginal walk L. In this section we will derive $E(U^k)$, $E(\otimes^2 U^k)$ and $\otimes^2 E(U^k)$, thus leading to the assembling of $\mathrm{Cov}(U^k)$.

4.2.1. *Derivation of $E(U^k)$.*

PROPOSITION 4.2. *If U is uniformly distributed in G, then $E(U^k) = a_0 ee' + a_1 I$, where a_0 and a_1 are determined by $n(a_0 + a_1) = E(tr\, U^k)$ and $n(n-1)a_0 = n - E(tr\, U^k)$.*

PROOF. Let $m_{ji} = \{g \in G; g^k(j) = i\}$, so that $|G|E(U^k)_{ij} = |m_{ji}|$. If $j \neq i$ then m_{ij} and any other m_{uv} with $u \neq v$ are isomorphic via $g \to t'gt$ where t is (any permutation) such that t(v)=j and t(u)=i; take g in m_{ij}, then $(t'gt)^k(v) = t'g^k t(v) = t'g^k(j) = t'(i) = u$, so that t'gt is in m_{uv}, and hence $|m_{ij}| = |m_{uv}|$ (off-diagonal homogeneity). Similarly, if j=i then m_{ii} and any other m_{ff} are isomorphic via $g \to t'gt$ where t is any permutation satisfying t(f)=i, and hence $|m_{ii}| = |m_{ff}|$ (diagonal homogeneity). Therefore, $E(U^k)$ has the form $a_0 ee' + a_1 I$, for constants a_0 and a_1, determined by taking the trace and overall sum on both sides of the equality $E(U^k) = a_0 ee' + a_1 I$. □

To determine the coefficients a_0 and a_1, we need the following: Given positive integers n and k let $D_{k,n}$ indicate the set of all positive integer divisors of k not greater than n, with $|D_{k,n}|$ elements in it. For computational purposes, note that $|D_{k,n}|$ may be expressed as $\sum_{r=1}^{n} I_{\{0\}}(k \mod (r))$, where $I_{\{0\}}$ is the indicator function with base the set $\{0\}$.

THEOREM 4.1. If U is uniformly distributed in G then $E(\text{tr } U^k) = |D_{k,n}|$.

PROOF. Let $c_{m,n}(U)$ indicate the random number of cycles of length m associated with a random permutation U in G, $m = 1, \ldots, n$. Then, noting that $\text{tr } g^k = \sum_{m \in D_{k,n}} m c_{m,n}(g)$, it follows that

$$E(\text{tr } U^k) = \sum_{m \in D_{k,n}} m E(c_{m,n}(U)).$$

However, because there are in average 1/m cycles of length m in a uniform random permutation in G, for all n, (that is $E(c_{m,n}(U)) = 1/m$) [see also Appendix B], the proposed result follows. □

From proposition 4.1, we obtain,

$$a_0 + a_1 = \frac{|D_{k,n}|}{n}, \quad a_0 = \frac{n - |D_{k,n}|}{n(n-1)}, \quad n a_0 + a_1 = 1. \tag{4.4}$$

The corresponding trace decomposition corresponding to Table 1 leads to $v_1 = (|D_{k,n}| - 1)/(n-1)$ and $v_0 = 1$, so that $\mathcal{F} = v_1 = (|D_{k,n}| - 1)/(n-1)$.

EXAMPLE 4.2. Calculations based on Theorem 4.1. In the following Tables 5 and 8, a permutation g is represented as $(1^{c_1}, 2^{c_2}, \ldots, n^{c_n})$, where $c_m \equiv c_{m,n}(g)$ is the multiplicity of cycles of length m in g, $m = 1, \ldots, n$. Cycles with null multiplicity are omitted in the notation shown in the tables. Table 5 shows the conjugacy classes $\mathcal{C}$, representatives of g^k, and the number $\gamma(g) = \sum c_{m,n}(g)$, for values of $k = 1, \ldots, 10$. We note that $\gamma(g)$ is the number of distinct orbits generated by g and its powers acting on $\{1, \ldots, n\}$ and that $|\mathcal{C}| = n!/\prod_{m=1}^{n} m^{c_m} c_m!$ (see also Appendix B). Table 6 shows values of tr g^k by conjugacy class $\mathcal{C}$, in S_5, for values of $k = 1, \ldots, 10$. Table 7 shows the probability distribution for $\tau_k = \text{tr } U^k$ and $E(\tau_k) = |D_{k,n}|$ in S_5, for values of $k = 1, \ldots, 10$ (also shown is $E(\tau_k^2)$). Table 8 shows values of $\tau_k = \text{tr } U^k = |D_{k,n}|$ in S_8 by conjugacy classes $\mathcal{C}$, for values of $k = 1, \ldots, 4$.

TABLE 5. Conjugacy classes $\mathcal{C}$, representatives of g^k, and number $\gamma(g)$ of $< g >$-orbits in S_5, for values of $k = 1, \ldots, 10$.

$\mathcal{C}$	$\gamma(g)$	$\lvert\mathcal{C}\rvert$	g	g^2	g^3	g^4	g^5	g^6	g^7	g^8	g^9	g^{10}
$\mathcal{C}_7$	1	24	5	5	5	5	1^5	5	5	5	5	1^5
$\mathcal{C}_6$	2	20	2, 3	$1^2, 3$	$2, 1^3$	$1^2, 3$	2, 3	1^5	2, 3	$1^2, 3$	$2, 1^3$	$1^2, 3$
$\mathcal{C}_5$	2	30	1, 4	$1, 2^2$	1, 4	1^5	1, 4	$1, 2^2$	1, 4	1^5	1, 4	$1, 2^2$
$\mathcal{C}_4$	3	20	$1^2, 3$	$1^2, 3$	1^5	$1^2, 3$	$1^2, 3$	1^5	$1^2, 3$	$1^2, 3$	1^5	$1^2, 3$
$\mathcal{C}_3$	3	15	$1, 2^2$	1^5	$1, 2^2$	1^5	$1, 2^2$	1^5	$1, 2^2$	1^5	$1, 2^2$	1^5
$\mathcal{C}_2$	4	10	$1^3, 2$	1^5	$1^3, 2$	1^5	$1^3, 2$	1^5	$1^3, 2$	1^5	$1^3, 2$	1^5
$\mathcal{C}_1$	5	1	1^5	1^5	1^5	1^5	1^5	1^5	1^5	1^5	1^5	1^5

TABLE 6. Values of $\tau_k = \text{tr } U^k$ by conjugacy class $\mathcal{C}$, in S_5, for values of $k = 1, \dots, 10$.

$\mathcal{C}$	$\gamma(g)$	$\lvert\mathcal{C}\rvert$	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9	τ_{10}
$\mathcal{C}_7$	1	24	0	0	0	0	5	0	0	0	0	5
$\mathcal{C}_6$	2	20	0	2	3	2	0	5	0	2	3	2
$\mathcal{C}_5$	2	30	1	1	1	5	1	1	1	5	1	1
$\mathcal{C}_4$	3	20	2	2	5	2	2	5	2	2	5	2
$\mathcal{C}_3$	3	15	1	5	1	5	1	5	1	5	1	5
$\mathcal{C}_2$	4	10	3	5	3	5	3	5	3	5	3	5
$\mathcal{C}_1$	5	1	5	5	5	5	5	5	5	5	5	5

TABLE 7. Probability distribution for $\tau_k = \text{tr } U^k$, $E(\tau_k) = |D_{k,n}|$ and $E(\tau_k^2)$ when U is uniformly distributed in S_5, for values of $k = 1, \dots, 10$.

τ	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9	τ_{10}
0	44	24	24	24	20	24	44	24	24	0
1	45	30	45	0	45	30	45	0	45	30
2	20	40	0	40	20	0	20	40	0	40
3	10	0	30	0	10	0	10	0	30	0
4	0	0	0	0	0	0	0	0	0	0
5	1	26	21	56	25	66	1	56	21	50
$E(\tau_k)$	1	2	2	3	2	3	1	3	2	3
$E(\tau_k^2)$	2	7	7	13	7	14	2	13	7	12

EXAMPLE 4.3. When n=2, U^k is either constant and equal to the identity when k is even, or U^k is either $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ or the identity, each with probability 1/2, when k is odd. Also note that $2 - |D_{k,2}| = k \mod 2$. From Proposition 4.1, it then follows that tr $\text{Cov}(U^k) = k \mod 2$ and $\text{Cov}(U^k) = \frac{k \mod 2}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$.

4.2.2. *Derivation of the diagonal blocks of $E(\otimes^2 U^k)$.* We start with the derivation of the expected trace of $U_{ii,k}U^k$. In what follows we write *marginal sum* to indicate both row sum and column sum without distinction. Let $G_{ij,k}$ indicate the subset of G consisting of those permutation matrices g such that the ij-entry of g^k is equal to 1, that is $G_{ij,k} = \{g \in G; e_i' g^k e_j = 1\}$.

PROPOSITION 4.3. *The size of $|G_{ij,k}|$ is $(n-1)!|D_{k,n}|$ when i=j and is $(n-2)!|D_{k,n}|$ when $i \neq j$.*

PROOF. In fact,

$$|G_{ii,k}| = |G|E(e_i' U^k e_i) = G\frac{|D_{k,n}|}{n} = |M_{n-1}||D_{k,n}| = (n-1)!|D_{k,n}|,$$

whereas, when $i \neq j$, $|G_{ij,k}| = |G|E(e_i' U^k e_j) = G\frac{n-|D_{k,n}|}{n(n-1)} = (n-2)!|D_{k,n}|$. □

TABLE 8. Values of $\tau_k = \text{tr } U^k$ in S_8 by conjugacy classes $\mathcal{C}$, and $E(\tau_k) = |D_{k,n}|$, for values of $k = 1, \ldots, 4$.

$\gamma(g)$	g	$\|\mathcal{C}\|$	τ_1	g^2	τ_2	g^3	τ_3	g^4	τ_4
1	8	5040	0	4^2	0	8	0	2^4	0
2	$1,7$	5760	1	$1,7$	1	$1,7$	1	$1,7$	1
2	$2,6$	3360	0	$1^2,3^2$	2	$2,2^3$	0	$1^2,3^2$	2
2	$3,5$	2688	0	$3,5$	0	$1^3,5$	3	$3,5$	0
2	4^2	1260	0	2^4	0	4^2	0	1^8	8
3	$1^2,6$	3360	2	$1^2,3^2$	2	$1^2,2^3$	2	$1^2,3^2$	2
3	$1,2,5$	4032	1	$1^3,5$	3	$1,2,5$	1	$1^3,5$	3
3	$1,3,4$	3360	1	$1,3,2^2$	1	$1^4,4$	4	$1^5,3$	5
3	$2^2,4$	1260	0	$1^4,2^2$	4	$2^2,4$	0	1^8	8
3	$2,3^2$	1120	0	$1^2,3^2$	2	$1^6,2$	6	$1^2,3^2$	2
4	$1^3,5$	1344	3	$1^3,5$	3	$1^3,5$	3	$1^3,5$	3
4	$1^2,2,4$	2520	2	$1^4,2^2$	4	$1^2,2,4$	2	1^8	8
4	$1^2,3^2$	1120	2	$1^2,3^2$	2	1^8	8	$1^2,3^2$	2
4	$1,2^2,3$	1680	1	$1^5,3$	5	$1^4,2^2$	5	$1^5,3$	5
4	2^4	105	0	1^8	8	2^4	0	1^8	8
5	$1^4,4$	420	4	$1^4,2^2$	4	$1^4,4$	4	1^8	8
5	$1^3,2,3$	1120	3	$1^5,3$	5	$1^6,2$	6	$1^5,3$	5
5	$1^2,2^3$	420	2	1^8	8	$1^2,2^3$	2	1^8	8
6	$1^5,3$	112	5	$1^5,3$	5	1^8	8	$1^5,3$	5
6	$1^4,2^2$	210	4	1^8	8	$1^4,2^2$	4	1^8	8
7	$1^6,2$	28	6	1^8	8	$1^6,2$	6	1^8	8
8	1^8	1	8	1^8	8	1^8	8	1^8	8
$E(\tau_k)$			1		2		2		3

Note that $G_{ii,k}$ is not a subgroup of G whenever $|D_{k,n}|$ is not an integer divisor of n. In particular, $G_{ii,1}$ is a subgroup of G.

PROPOSITION 4.4. *If U is uniformly distributed in G, then $E(U_{ii,k}$ tr $U^k) = \frac{|D_{k,n}|}{n} \sum_{m \in D_{k,n}} m\widehat{C}_{m,k}$, where $\widehat{C}_{m,k}$ is the average number of m-cycles in $G_{ii,k}$.*

PROOF.

$$\begin{aligned}
E(U_{ii,k}\text{tr } U^k) &= \sum_G e_i' g^k e_i \text{tr } g^k/|G| = \sum_{G_{ii,k}} \text{tr } g^k/|G| \\
&= \frac{|G_{ii,k}|}{|G|} \sum_{G_{ii,k}} \text{tr } g^k/|G_{ii,k}| = \frac{|D_{k,n}|}{n} \sum_{G_{ii,k}} \text{tr } g^k/|G_{ii,k}| \\
&= \frac{|D_{k,n}|}{n} \sum_{G_{ii,k}} \sum_{m \in D_{k,n}} m c_{m,n}(g)/|G_{ii,k}| \\
&= \frac{|D_{k,n}|}{n} \sum_{m \in D_{k,n}} m \sum_{G_{ii,k}} c_{m,n}(g)/|G_{ii,k}| \\
&= \frac{|D_{k,n}|}{n} \sum_{m \in D_{k,n}} m\widehat{C}_{m,k},
\end{aligned}$$

where $\widehat{C}_{m,k} = \sum_{G_{ii,k}} c_{m,n}(g)/|G_{ii,k}|, m = 1, \ldots, n$ is the average number of m-cycles in $G_{ii,k}$. □

To obtain the main diagonal of $E(U_{ii,k}U^k)$ we start with the (ii,ii) entry. From (4.4),

$$E(U_{ii,k}^2) = E(U_{ii,k}) = \frac{|D_{k,n}|}{n}. \tag{4.5}$$

The n-1 remaining entries of the main diagonal are

$$E(U_{ii,k}U_{ss,k}) = \frac{E(U_{ii,k} \text{ tr } U^k) - |D_{k,n}|/n}{n-1}, \quad s \neq i. \tag{4.6}$$

Proposition 4.5. *If* $\mathbf{U}$ *is uniformly distributed in G, then: (a) The entries* $i \neq s$ *of column t, and the entries* $j \neq t$ *of row s of* $E(U_{ij,k}U^k)$ *are equal to zero; (b) The marginal sums of* $E(U_{ij,k}U^k)$ *are constant and equal to* $\frac{|D_{k,n}|}{n}$ *for i=j, and* $\frac{n-|D_{k,n}|}{(n-1)}$ *for* $i \neq j$.

Proof. Part (a) follows from the fact that each row (column) of U^k has exactly one entry equal to 1 and the remaining entries equal to 0; For (b), note that $E(\sum_s U_{ij,k}U_{st,k}) = E(U_{ij,k}\sum_s U_{st,k}) = E(U_{ij,k})$. Applying Propositions 4.2 and 4.1, the result follows (similarly for column sums). □

Proposition 4.5, and equations (4.5) and (4.6) directly determine all the entries at column t=j and row s=i. Any other row (say row $s \neq i$) has the diagonal entry (s,s) given by (4.6), the (s,i) entry equal to zero and (for $n > 2$) the remaining n-2 entries given by

$$E(U_{ii,k}U_{st,k}) = \frac{|D_{k,n}|/n - E(U_{ii,k}U_{ss,k})}{n-2}, \quad t \neq i, s. \tag{4.7}$$

Example 4.4. Evaluation of $E(U_{11,k}U^k \text{ tr } U^k)$ for n=3 and k=1,...,6. Table 9 summarizes the computational steps of Proposition 4.4. Consider the case k=2, to illustrate. First we evaluate $|G_{11,k}| = (n-1)!|D_{k,n}| = 4$. The corresponding permutations are marked with 1 on the box to the left of column $|G_{11,k}|$. Next, $\hat{C}_{1,k}$, is obtained by summing $C_1(g)$ over those columns corresponding to permutations g marked with 1 (the first four when k=2) and dividing by $|G_{11,k}|$. This leads to $\widehat{C}_{1,2} = 6/4$. Next, compute $W = \sum_{m \in D_{k,n}} m\widehat{C}_{m,k}$, or, is this case, $\sum_{m=1,2} m\widehat{C}_{m,2} = 6/4 + 2 \times 3/4 = 12/4$. Finally, $E(U_{11,2}\text{tr } U^2) = W|D_{2,3}|/3 = 2$. The derivations for other values of k are obtained similarly.

4.2.3. *Derivation of the off-diagonal blocks of* $E(\otimes^2 U^k)$. Let [i j] indicate the permutation matrix representing the transposition (ij). Direct verification shows that $U_{ii,k}U^k \sim U_{jj,k}[i\ j]U^k[i\ j]'$, and that, for $i \neq j$, $U_{ij,k}U^k \sim U_{im,k}[j\ m]U^k[j\ m]'$ (in each of the cases (a) $s, t \in \{j, m\}$, (b) $s \in \{j, m\}$ and $t \notin \{j, m\}$, (c) $s \notin \{j, m\}$ and $t \in \{j, m\}$, and (d) $s, t \notin \{j, m\}$, it holds that $U_{ij,k}e_s'U^ke_t \sim U_{im,k}e_s'[j\ m]U^k[j\ m]'e_t$).

Proposition 4.6. *The following equalities hold:*

(1) $E(U_{ii,k} tr\ U^k) = \frac{1}{n}E(tr^2\ U^k)$,
(2) $E(U_{ij,k} tr\ U^k) = \frac{1}{n-1}E(tr\ U^k) - \frac{1}{n(n-1)}E(tr^2\ U^k), \quad i \neq j$.

TABLE 9. Values of $|G_{11,k}|$, $\hat{C}_{m,k}$, $W = \sum_{m\in D_{k,n}} m\widehat{C}_{m,k}$ and $E(U_{11,k}\text{tr } U^k) = W|D_{k,n}|/n$, for n=3 and values of k=1,...,6.

k	$D_{k,n}$	1	(12)	(13)	(23)	(123)	(132)
1	$\{1\}$	1	0	0	1	0	0
2	$\{1,2\}$	1	1	1	1	0	0
3	$\{1,3\}$	1	0	0	1	1	1
4	$\{1,2\}$	1	1	1	1	0	0
5	$\{1\}$	1	0	0	1	0	0
6	$\{1,2,3\}$	1	1	1	1	1	1
	$C_1(g)$	3	1	1	1	0	0
	$C_2(g)$	0	1	1	1	0	0
	$C_3(g)$	0	0	0	0	1	1

k	$\|G_{11,k}\|$	$\hat{C}_{1,k}$	$\hat{C}_{2,k}$	$\widehat{C}_{3,k}$	W	$E(U_{11,k}\text{tr } U^k)$
1	2	4/2	1/2	0/2	4/2	2/3
2	4	6/4	3/4	0/4	12/4	2
3	4	4/4	1/4	2/4	10/4	5/3
4	4	6/4	3/4	0/4	12/4	2
5	2	4/2	1/2	0/2	4/2	2/3
6	6	6/6	3/6	2/6	18/3	3

PROOF. Write (the power notation is irrelevant and is omitted here), for part (1), $(\text{tr } U)U = U_{11}U + U_{22}U + \ldots + U_{nn}U$. From Proposition 4.6, $(\text{tr } U)U \sim U_{11}U + U_{22}[1\ 2]U[1\ 2]' + \ldots + U_{nn}[1\ n]U[1\ n]'$, so that, $(\text{tr } U)^2 \sim U_{11}\text{tr } U + U_{22}\text{tr } U + \ldots + U_{nn}\text{tr } U \sim nU_{11}\text{tr } U$. Taking the expectation on both sides, the result follows. For part (2), again from Proposition 4.6, we write, for $m \neq i$,

$$\sum_{j\neq i} U_{im}[j\ m]U[j\ m]' \sim \sum_{j\neq i} U_{ij}U = (1 - U_{ii})U, \tag{4.8}$$

so that $(n-1)U_{im}\text{tr } U \sim \text{tr } U - \text{tr } U_{ii}U$. Taking expectations on both sides, and applying part (a), the result follows. □

EXAMPLE 4.5. Suppose n=4. The derivation of $Z_{12} = E(U_{12,k}U^k)$ is obtained as follows. First derive $M = E(1 - U_{11,k}U^k)$, following Example 4.4. Then apply Propositions 4.5 (b) and 4.6 (b), along with the estimating equations $M = Z_{12} + [2\ 3]Z_{12}[2\ 3]' + [2\ 4]Z_{12}[2\ 4]'$, to determine the components of Z_{12}. The remaining blocks on the first row are obtaining by conjugating Z_{12} (e.g., $Z_{13} = E(U_{13,k}U^k) = [2\ 3]Z_{12}[2\ 3]'$). The off-diagonal blocks on the remaining rows are complement-symmetric: $Z_{ij} = Z_{n-j,n-i}$ (e.g., $Z_{23} = Z_{12}$).

EXAMPLE 4.6. The following are the transition probability blocks,

$$P = E(C_1, [C_1', C_2', C_3'])$$

, and corresponding covariance matrices, $C = \text{Cov}(C_1, [C_1', C_2', C_3'])$ for n=3 and k=1,...,6, and the complete 9×9 matrices P and C for k=4 (lexicographic notation).

k=1

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 2 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 & 1 & 0 \end{array}\right],$$

$$C = \frac{1}{18}\left[\begin{array}{rrr|rrr|rrr} 4 & -2 & -2 & -2 & 1 & 1 & -2 & 1 & 1 \\ -2 & 4 & -2 & 1 & -2 & 1 & 1 & -2 & 1 \\ -2 & -2 & 4 & 1 & 1 & -2 & 1 & 1 & -2 \end{array}\right]$$

k=2

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right],$$

$$C = \frac{1}{36}\left[\begin{array}{rrr|rrr|rrr} 8 & -4 & -4 & -4 & 8 & -4 & -4 & -4 & 8 \\ -4 & 5 & -1 & -1 & -4 & 5 & 5 & -1 & -4 \\ -4 & -1 & 5 & 5 & -4 & -1 & -1 & 5 & -4 \end{array}\right]$$

k=3

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 4 & 0 & 0 & 0 & 3 & 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \end{array}\right],$$

$$C = \frac{1}{36}\left[\begin{array}{rrr|rrr|rrr} 8 & -4 & -4 & -4 & 2 & 2 & -4 & 2 & 2 \\ -4 & 5 & -1 & 5 & -4 & -1 & -1 & -1 & 2 \\ -4 & -1 & 5 & -1 & 2 & -1 & 5 & -1 & -4 \end{array}\right]$$

k=4

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right],$$

$$C = \frac{1}{36}\left[\begin{array}{rrr|rrr|rrr} 8 & -4 & -4 & -4 & 8 & -4 & -4 & -4 & 8 \\ -4 & 5 & -1 & -1 & -4 & 5 & 5 & -1 & -4 \\ -4 & -1 & 5 & 5 & -4 & -1 & -1 & 5 & -4 \end{array}\right]$$

k=5

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 2 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 & 1 & 0 \end{array}\right],$$

$$C = \frac{1}{18}\left[\begin{array}{rrr|rrr|rrr} 4 & -2 & -2 & -2 & 1 & 1 & -2 & 1 & 1 \\ -2 & 4 & -2 & 1 & -2 & 1 & 1 & -2 & 1 \\ -2 & -2 & 4 & 1 & 1 & -2 & 1 & 1 & -2 \end{array}\right]$$

k=6

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 6 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

$$C = \frac{1}{18}\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Complete transition probability and covariance matrices for k=4, in lexicographic notation.

$$P = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 \end{array}\right],$$

$$\mathrm{Cov}(U^4) = \frac{1}{36}\left[\begin{array}{rrr|rrr|rrr} 8 & -4 & -4 & -4 & 8 & -4 & -4 & -4 & 8 \\ -4 & 5 & -1 & -1 & -4 & 5 & 5 & -1 & -4 \\ -4 & -1 & 5 & 5 & -4 & -1 & -1 & 5 & -4 \\ \hline -4 & -1 & 5 & 5 & -4 & -1 & -1 & 5 & -4 \\ 8 & -4 & -4 & -4 & 8 & -4 & -4 & -4 & 8 \\ -4 & 5 & -1 & -1 & -4 & 5 & 5 & -1 & -4 \\ \hline -4 & 5 & -1 & -1 & -4 & 5 & 5 & -1 & -4 \\ -4 & -1 & 5 & 5 & -4 & -1 & -1 & 5 & -4 \\ 8 & -4 & -4 & -4 & 8 & -4 & -4 & -4 & 8 \end{array}\right]$$

Complete transition probability and covariance matrices for k=4, in $\otimes$-notation.

$$P_{\otimes} = \frac{1}{6}\left[\begin{array}{ccc|ccc|ccc} 4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 4 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 4 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 4 & 1 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 4 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 4 \end{array}\right]$$

$$\mathrm{Cov}(U^4)_\otimes = \frac{1}{36}\left[\begin{array}{ccc|ccc|ccc} 8 & -4 & -4 & -4 & 5 & -1 & -4 & -1 & 5 \\ -4 & 8 & -4 & -1 & -4 & 5 & 5 & -4 & -1 \\ -4 & -4 & 8 & 5 & -1 & -4 & -1 & 5 & -4 \\ \hline -4 & -1 & 5 & 8 & -4 & -4 & -4 & 5 & -1 \\ 5 & -4 & -1 & -4 & 8 & -4 & -1 & -4 & 5 \\ -1 & 5 & -4 & -4 & -4 & 8 & 5 & -1 & -4 \\ \hline -4 & 5 & -1 & -4 & -1 & 5 & 8 & -4 & -4 \\ -1 & -4 & 5 & 5 & -4 & -1 & -4 & 8 & -4 \\ 5 & -1 & -4 & -1 & 5 & -4 & -4 & -4 & 8 \end{array}\right]$$

5. Concluding remarks

The covariance structure of U^k is naturally related to the cyclic structure of the symmetric group. Here we will outline that connection from two perspectives: one based on simple combinatoric arguments and the other based on elementary character theory. First, note that Propositions 4.4 and 4.6 show that in the uniform case,

$$E(tr^2\ U^k) = |D_{k,n}| \sum_{m \in D_{k,n}} m\widehat{C}_{m,k}, \tag{5.1}$$

where $\widehat{C}_{m,k}$ is the average number of m-cycles in $G_{ii,k} = \{g \in G; e_i' g^k e_i = 1\}$. In particular, when k=1, one shows that

$$\sum_{g \in G_{ii,1}} \mathrm{tr}\ g = (n-1)! + \sum_{h \in S_{n-1}} \mathrm{tr}\ h. \tag{5.2}$$

Dividing both sides of equation (5.2) by $|G_{ii,1}| = (n-1)!$ (see Proposition 4.3), it follows that $\widehat{C}_{1,1} = 1 + E(\mathrm{tr}\ Z)$, with Z uniformly in S_{n-1}. Because $E(\mathrm{tr}\ Z) = 1$, we obtain $E(tr^2\ U) = \widehat{C}_{1,1} = 2$. As outlined below (Appendix A, Comment 3), one obtains the same result using Burnside's Lemma. The question of interest here is the determination of $E(tr^2\ U^k)$ for arbitrary k, when U is uniform in G. Because

$$\mathrm{tr}\ U^k = \sum_{m \in D_{k,n}} m c_{m,n}(U),$$

it follows that $E(tr^2\ U^k)$ is completely determined by the joint covariance structure of

$$c_n'(U) = (c_{1,n}(U), \ldots, c_{n,n}(U)).$$

In fact, it is only necessary to determine the matrix $E(c_n(U)c_n'(U))$ of average cross-products.

5.1. The combinatorial approach. The joint probability law of $c_n(U)$ has the support the set K of points $c' = (c_1, \ldots, c_n)$ with non-negative integer coordinates satisfying $\sum_{m=1}^n mc_m = n$. Clearly, $|K|$ equals the number of conjugacy classes in G and the law of $c_n(U)$ assigns the same probability to members within the same conjugacy class $|\mathcal{C}|$. Therefore $P(c_n(U) = c) = |\mathcal{C}|/n!$ for each $c \in K$, where $\mathcal{C}$ is the conjugacy class with representative $(1^{c_1}, \ldots, m^{c_m})$. Standard combinatoric arguments (e.g., [**Tak84**]) show that there are exactly $n!/\prod_K m^{c_m} c_m!$, so that, equivalently, $P(c_n(U) = c) = 1/\prod_K m^{c_m} c_m!$. Given the law of $c_n(U)$, the matrix $E(c_n(U)c_n'(U))$ is then obtained by its own definition. The following illustrates the case n=3 and k=2.

EXAMPLE 5.1. Derivation of $E(tr^2\ U^2)$ in S_3. There are $|K| = 3$ conjugacy classes in S_3 and the points in the support set K are $c = (0,0,1)$, with probability 1/3, $c = (1,1,0)$ with probability 1/2 and $c = (3,0,0)$ with probability 1/6. The resulting marginals $c_{m,3}(U) \equiv c_m(U), m = 1,2,3$ have support the set $\{0,1,2,3\}$, with corresponding probabilities $(1/3, 1/2, 0, 1/6)$ for $c_1(U)$, $(1/2, 1/2, 0, 0)$ for $c_2(U)$ and $(2/3, 1/3, 0, 0)$ for $c_3(U)$. Note that $E(c_m(U)) = 1/m$, whereas $E(c_1^2(U)) = E(tr^2\ U) = 2$, as discussed above. In addition, $E(c_2^2(U)) = 1/2$ and $E(c_3^2(U)) = 1/3$. Similarly, the joint bivariate marginals are determined. The law of $(c_1(U), c_2(U))$ assigns probabilities 1/3,1/2 and 1/6 to the points (0,0),(1,1) and (3,0), respectively, so that $E(c_1(U)c_2(U)) = 1/2$, whereas $(c_1(U), c_3(U))$ takes probabilities 1/3,1/2 and 1/6 at (0,1),(1,0) and (3,0), respectively, with $E(c_1(U)c_3(U)) = 0$. The law of $(c_1(U), c_2(U))$ assigns probabilities 1/6,1/3 and 1/2 to the points (0,0),(0,1) and (1,0), respectively, so that $E(c_2(U)c_3(U)) = 0$. In summary, we obtain,

$$E(c_3(U)c_3'(U)) = \begin{bmatrix} 2 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}, \quad \mathrm{Cov}(c_3(U)) = \begin{bmatrix} 1 & 0 & -\frac{1}{3} \\ 0 & \frac{1}{4} & -\frac{1}{6} \\ -\frac{1}{3} & -\frac{1}{6} & \frac{1}{9} \end{bmatrix}.$$

Then, it follows that

$$\begin{aligned} E(tr^2\ U^2) &= E(c_{1,3}(U) + 2c_{2,3}(U))^2 \\ &= E(c_{1,3}^2(U)) + 4E(c_{1,3}(U)c_{2,3}(U)) + 4E(c_{2,3}^2(U)) = 6. \end{aligned}$$

5.2. The character theory approach. Let Θ indicate the $|K| \times |K|$ matrix of character values of G, with entries $\theta_{ij} = \chi_i(g_j)$, where g_j is a representative in the conjugacy class $\mathcal{C}_j$, $j = 1, \ldots, |K|$. Usually, χ_1 is the character of the trivial representation. Let T indicate the $n \times |K|$ matrix with entries

$$T_{mj} = \frac{|\mathcal{C}_j|}{n!} c_{m,n}(g_j), \quad j = 1, \ldots, |K|, \quad m = 1, \ldots, n.$$

One then shows that the components of m-th column of the matrix $\Theta T'$ are exactly the Fourier coefficients

$$< c_{m,n}, \chi_j >= \frac{1}{n!} \sum_G c_{m,n}(g) \bar{\chi}_j(g) = \frac{|\mathcal{C}_j|}{n!} \sum_{l=1}^{|K|} c_{m,n}(g_l) \bar{\chi}_j(g_l),$$

of the class function $c_{m,n}(g)$ relative to character χ_j, $j = 1, \ldots, |K|$. Moreover,

$$E(c_n(U)c_n'(U)) = T\Theta'\Theta T', \tag{5.3}$$

and $E(c_n(U)) = T\Theta' e_1$, so that $\mathrm{Cov}(c_n(U)) = T\Theta'\Theta T' - T\Theta' e_1 e_1' \Theta T'$. The following example illustrates the result:

EXAMPLE 5.2. We consider again the case n=3, as in Example 5.1 above. We start with the characters matrix Θ, where we define the conjugacy class $\mathcal{C}_1$ with representative 1^3, $\mathcal{C}_2$ with representative 2^1 and $\mathcal{C}_3$ with representative 3^1, so that

$$\Theta = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & -1 \end{bmatrix}.$$

From Example 5.1 we obtain the matrix T, as defined (note that the columns of T are indexed by the same labels defined by Θ),

$$T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

From (5.3), we obtain

$$E(c_n(U)c_n'(U)) = T\Theta'\Theta T' = T \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} T' = \begin{bmatrix} 2 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix},$$

which is as derived in Example 5.1.

Appendix A. The (n-1)-dimensional irreducible representation of G.

The existence and construction of the representation is outlined in [**Ser77**, p.17] and also in [**Dia88**, pp.6,134]. Let ρ indicate the permutation representation on the symmetric group (G) acting on a set X with n objects and let Y indicate the product of p copies of X, $p = 1, \dots, n$. Let also χ indicate the character of ρ.

(1) The character of ρ acting on Y is χ^p;
(2) When p=1 or p=2, Y has p distinct orbits (G is n-transitive, e.g., [**Rob93**, Ch.7]). When p=3, there are 5 distinct representatives, (1,1,1), (1,1,f) , (1,f,1), (f,1,1), with $f \neq 1$ and (t,f,g), with $t \neq f, f \neq g, t \neq g$;
(3) In general, |G-orbits in Y| $= \sum_G \chi^p(g)/|G|$. This is essentially Burnside's Lemma [see [**Dia88**, p.134]]. This also gives all moments of the trace function of random permutations, and |G-orbits in Y| gives the first n moments of a Poisson (1) distribution. In particular, $E(\text{tr } U) = 1$ and $E(tr^2\ U) = 2$;
(4) Let 1 indicate the identity representation, and θ its complementary subspace. Let the corresponding characters be χ_1 and χ_θ. Burnside Lemma implies that $1 = (\chi|\chi_1) = \sum \chi(g)/|G|$, so that 1 is contained in ρ exactly once. Write $\rho = 1 + \theta$;
(5) From 2 and 3 above, $2 = (\chi^2|1) = (\chi|\chi) = (1+\theta|1+\theta) = 1 + (\theta|\theta)$, so that $(\theta|\theta) = 1$ and hence θ is irreducible;
(6) The homotheties in V and W are the corresponding projections, given by $\sum_G \chi_1(g)\rho(g)\text{dim V}/|G|$ and $\sum_G \chi_\theta(g)\rho(g)\text{dim W}/|G|$, respectively. The homothety ratios are tr H_{11}/dim V and tr H_{22}/dim W, where H_{11} and H_{22} are the block partition of H corresponding to the dimensions of V and W, respectively. In the present case, the ratios are h_{11} and $(h_{22} + \ldots + h_{nn})/(n-1)$. More generally, Schur's Lemma shows that

$$\sum_G (\theta_1(g) \vdots \theta_2(g)) H (\theta_1(g') \vdots \theta_2(g'))/|G| = diag\ (\frac{\text{tr } H_{11}}{d_1} I_1, \frac{\text{tr } H_{22}}{d_2} I_2),$$

for any two irreducible non-isomorphic representations θ_1 and θ_2 of dimensions d_1 and d_2 (and H of dimension $d_1 + d_2$).

Appendix B. The average number of cycles and m-cycles

Let $c_{m,n}(U)$ indicate the number of m-cycles in a random uniform permutation U in G, so that $\gamma_n(U) = \sum_{m=1}^n c_{m,n}(U)$ is the number of cycles in U. The average number $\bar{c}_{m,n}$ of cycles of length m (m-cycles) in G is 1/m, for $m = 1, \ldots, n$ and all n. This follows from Burnside Lemma (e.g., [**DM96**, p.25]). The number of permutations in G with exactly k cycles is the Stirling number S(n,k) [e.g., [**Tak84**]], so that (i) $\gamma_n(U) = k$ with probability S(n,k)/n!, k=1,...,n, (ii) the mean number $\bar{\gamma}_n = E(\gamma_n(U))$ of cycles in G is

$$\bar{\gamma}_n = E(\gamma_n(U)) = \sum_{k=1}^n k\frac{S(n,k)}{n!} = \sum_{m=1}^n E(c_{m,n}(U)) = \sum_{m=1}^n \bar{c}_{m,n},$$

and (iii) $\bar{c}_{m,n}$ is the average number of m-cycles in G.

PROPOSITION B.1. *The mean number of cycles in G is* $\bar{\gamma}_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$.

PROOF. This follows directly from the fact that $\bar{c}_{m,n} = 1/m$ for $m = 1, \ldots, n$ and all n. A direct proof is based on the recursive relation $S(n+1,k) - nS(n,k) = S(n,k-1)$ for Stirling numbers. Simple calculation then shows that $\bar{\gamma}_{n+1} = \bar{\gamma}_n + \frac{1}{n+1}$, from which the result follows. □

Because $\gamma_n(U)$ is also the number of orbits generated by the multiplicative action of G on $\{1, \ldots, n\}$, and $\gamma_n(U)$ is constant over conjugacy classes, we also obtain

$$\sum_j \gamma_j \frac{|\mathcal{C}_j|}{n!} = \bar{\gamma}_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n},$$

where the sum is indexed by the conjugacy classes and γ_j is the common value of $\gamma_n(U)$ for U in $\mathcal{C}_j$.

References

[BE73] M. J. Buckley and G. K. Eagleson, *Assessing large sets of rank correlations*, Biometrika **73** (1973), no. 1, 151–7.

[Dan62] H. E. Daniels, *Processes generating permutation expansions*, Biometrika **49** (1962), 139–149.

[DF99] P. Diaconis and D. Freedman, *Iterated random functions*, SIAM Reviw **41** (1999), no. 1, 45–76.

[Dia88] Persi Diaconis, *Group representation in probability and statistics*, IMS, Hayward, California, 1988.

[DM96] J. D. Dixon and B. Mortimer, *Permutation groups*, Springer, New York, NY, 1996.

[KW89] S. Klyce and S. Wilson, *Methods of analysis of corneal topography*, Refractive and Corneal Surgery **5** (1989), 368–371.

[LV99] H. Lee and M. A. G. Viana, *The joint covariance structure of ordered symmetrically dependent observations and their concomitants of order statistics*, Statistics and Probability Letters **43** (1999), 411–414.

[Nai82] M. A. Naimark, *Theory of group representations*, Springer-Verlag, New York, NY, 1982.

[OV95] I. Olkin and M. A. G. Viana, *Correlation analysis of extreme observations from a multivariate normal distribution*, Journal of the American Statistical Association **90** (1995), 1373–1379.

[PGL+88] V. Perkel, M. Gail, J. Lubin, D. Pee, R. Weinstein, E. Shore-Freedman, and A. Schneider, *Radiation-induced thyroid neoplasms: Evidence for familial susceptibility factors*, Journal of Clinical Endocrinology and Metabolism **66** (1988), no. 6, 1316–22.

[Rob93] Derek J.S. Robinson, *A course in the theory of groups*, Springer-Verlag, New York, 1993.

[Ser77] Jean-Pierre Serre, *Linear representations of finite groups*, Springer-Verlag, New York, 1977.

[Tak84] L. Takács, *Combinatorics*, Handbook of Statistics (New York, NY) (P. R. Krishnaiah and P. K. Sen, eds.), vol. 4, North-Holland, New York, NY, 1984, pp. 123–173.

[Via98] M. A. G. Viana, *Linear combinations of ordered symmetric observations with applications to visual acuity*, Order Statistics: Applications (Amsterdam) (N. Balakrishnan and C. R. Rao, eds.), vol. 17, Elsevier, Amsterdam, 1998, pp. 513–24.

[VO97] M. A. G. Viana and I. Olkin, *Correlation analysis of ordered observations from a block-equicorrelated multivariate normal distribution*, Advances in Statistical Decision Theory and Applications (Boston) (S. Panchapakesan and N. Balakrishnan, eds.), Birkhauser, Boston, 1997, pp. 305–322.

[VOM93] M. A. G. Viana, I. Olkin, and T. McMahon, *Multivariate assessment of computer analyzed corneal topographers*, Journal of the Optical Society of America - A **10** (1993), no. 8, 1826–1834.

Eye Research Institute, Division of Biostatistics, and Department of Mathematics, Statistics and Computer Science, The University of Illinois at Chicago (m/c 648), 1855 W. Taylor Street, Chicago, IL 60612.

E-mail address: viana@uic.edu