FREQUENCY SQUARES OF ORDERS 7 AND 8

VEDRAN KRČADINAC

Abstract. We enumerate frequency squares of orders 7 and 8 and the corresponding isotopism classes. An error in a previous paper by L.J. Brant and G.L. Mullen is discovered and partially corrected.

1. Introduction

Frequency squares are a generalization of Latin squares. Let $\lambda_1, \ldots, \lambda_s$ be positive integers and $n = \lambda_1 + \ldots + \lambda_s$. A *frequency square* of order n with frequency vector $\lambda = (\lambda_1, \ldots, \lambda_s)$ is an $n \times n$ matrix over $\{1, 2, \ldots, s\}$ such that the number i appears exactly λ_i times in each row and column, for $i = 1, \ldots, s$. Latin squares are frequency squares with $\lambda = (1, \ldots, 1)$. If the entries in the first row and column appear in natural order, the frequency square is said to be *reduced*. The set of all frequency squares will be denoted $F(n; \lambda)$ and the set of reduced frequency squares $f(n; \lambda)$. Two frequency squares are *isotopic* provided they are equivalent under rearrangements of rows, columns and entries.

Enumeration of frequency squares and isotopism classes is, in general, hard. There are several published formulae for the number of Latin squares (e.g. [8], [9] and [11]) but they are computationally infeasible. Reduced Latin squares have recently been enumerated by computer for $n = 11$ [8] and isotopism classes of Latin squares for $n = 9, 10$ [7]. The last published enumeration of frequency squares seems to be Brant and Mullen [1]. They determined numbers of frequency squares for $n \leq 6$ and with some exceptions also the corresponding isotopism classes. In this work an orderly classification algorithm is used to constructively enumerate isotopism classes of frequency squares up to $n = 8$. An independent enumeration of reduced frequency squares is also performed and the results are shown to be consistent with each other. The numbers imply that [1, Theorem 3] is wrong. Errors in the proof are pointed out and partially corrected, thereby proving [1, Corollary 4].

Date: March 30, 2005.
2000 *Mathematics Subject Classification.* 05B15.
Key words and phrases. Frequency square, enumeration, classification.

2. Isotopism classes

Orderly classification algorithms were pioneered by R.C. Read [10] and I.A. Faradžev [4]. We first give a general description of the method in the spirit of [6]. Let X be a set of combinatorial objects with a distinguished object $O \in X$. Let $p : X \to X$ be a nilpotent function, such that $p^n(X) = \{O\}$ for some $n \in \mathbb{N}$. Thus, X is a tree rooted at O and p is the parent function. The *order* of an object $A \in X$ is the level at which A appears in the tree: $\operatorname{ord} A = \min\{i \in \mathbb{N}_0 \,|\, p^i(A) = O\}$. Furthermore, suppose a group G acts on X in an order-preserving manner, i.e. $\operatorname{ord}(gA) = \operatorname{ord} A$ for all $g \in G$ and $A \in X$. The equivalence relation induced by the action of G will be denoted $\cong$. The goal is to construct one representative from each equivalence class, usually on a specified level of the tree.

We need one more ingredient to formulate the classification algorithm: a *canonical labeling map*. This is a function $c : X \to X$ with the properties $c(A) \cong A$ and $c(gA) = c(A)$, for all $A \in X$ and $g \in G$. The fixed points of c are called *canonical objects*. The canonical objects constitute a set of class representatives and thus a solution of the classification problem. However, the total number of objects is usually very large and it is not possible to examine them all in order to find the canonical ones. The key idea is to construct canonical objects recursively, by extending canonical objects on previous levels of the tree. This will be possible if parents of canonical objects are also canonical, i.e. if $c(A) = A$ implies $c(p(A)) = p(A)$. Now we have everything in place for a precise description of the algorithm.

```
scan(A, n)
[ if ord A = n then print A
[ else for all B ∈ p^{-1}(A) do
[   [ if c(B) = B then scan(B, n)
```

The call $\operatorname{scan}(O, n)$ will print all canonical objects of order n.

We want to construct a list of isotopism class representatives of frequency squares in $F(n;\lambda)$. As partial objects we take *frequency rectangles*, i.e. $r \times n$ matrices over $\{1, \ldots, s\}$ such that the number i appears at most λ_i times in every row and column. The set of all such matrices will be denoted $F(r, n; \lambda)$; obviously $F(n, n; \lambda) = F(n; \lambda)$. Take $X = \cup_{r=0}^{n} F(r, n; \lambda)$ and deletion of the last matrix row as the parent function p. Then the order of a frequency rectangle $A \in X$ is its number of rows. The group $G_r = S_r \times S_n \times (S_s)_\lambda$ acts on $F(r, n; \lambda)$ by permuting rows, columns and entries. Here S_r and S_n are symmetric groups and $(S_s)_\lambda$ is the group of s-permutations leaving the frequency vector λ invariant. The direct product $G = G_0 \times$

$\ldots \times G_n$ acts on X in an order-preserving way and induces the isotopism relation.

Next, we introduce a total ordering relation on the set of objects. Frequency rectangles are compared lexicographically, as vectors obtained by concatenating the rows. A canonical labeling map can be defined by $c(A) = \min\{gA \,|\, g \in G\}$ ($c(A)$ is the minimal frequency rectangle isotopic to A). Obviously, if A is the minimal rectangle in its isotopism class, the rectangle obtained by deleting the last row shares this property. Hence, parents of canonical frequency rectangles are also canonical and the classification algorithm can be applied in this setting.

The algorithm was implemented in the programming language C and run on a cluster of Linux workstations. For $n \leq 6$ our results agree with those of [1]. The missing numbers for $n = 6$ and numbers of isotopism classes of order $n = 7, 8$ are reported in Table 1. Complete lists of representatives can be accessed through the author's web page `http://www.math.hr/~krcko`, as well as the computer programs used for the calculations.

n	λ	No. cl.	n	λ	No. classes
6	$(2,2,2)$	46	8	$(5,3)$	51
	$(2,2,1,1)$	106		$(5,2,1)$	624
	$(2,1,1,1,1)$	56		$(5,1,1,1)$	370
7	$(5,2)$	4		$(4,4)$	156
	$(5,1,1)$	4		$(4,3,1)$	19 041
	$(4,3)$	16		$(4,2,2)$	112 043
	$(4,2,1)$	92		$(4,2,1,1)$	347 263
	$(4,1,1,1)$	56		$(4,1,1,1,1)$	93 561
	$(3,3,1)$	226		$(3,3,2)$	766 361
	$(3,2,2)$	1 939		$(3,3,1,1)$	1 211 710
	$(3,2,1,1)$	5 300		$(3,2,2,1)$	27 865 024
	$(3,1,1,1,1)$	1 398		$(3,2,1,1,1)$	29 632 348
	$(2,2,2,1)$	15 269		$(3,1,1,1,1,1)$	4 735 238
	$(2,2,1,1,1)$	22 813		$(2,2,2,2)$	26 983 466
	$(2,1,1,1,1,1)$	6 941		$(2,2,2,1,1)$	171 710 120
	$(1,1,1,1,1,1,1)$	564		$(2,2,1,1,1,1)$	137 000 435
8	$(6,2)$	7		$(2,1,1,1,1,1,1)$	29 163 047
	$(6,1,1)$	7		$(1,1,1,1,1,1,1,1)$	1 676 267

TABLE 1. Number of isotopism classes of frequency squares.

3. Reduced frequency squares

Enumeration and classification of Latin squares is a subject that produced many errors in published sources. One example is the value 563 as the number of isotopism classes of 7×7 Latin squares (e.g. in [2] and [3]). Table 1 comprises the correct value, 564. For a history of this and other errors see [7].

The total number of distinct frequency squares will be denoted by $|F(n;\lambda)|$ and $|f(n;\lambda)|$ will denote the total number of distinct reduced frequency squares. Any classification of as many objects as there are frequency squares of order 8 necessarily includes many reductions. There is a real danger of missing some classes due to conceptual and programming errors. We felt an independent verification of the numbers in Table 1 was desirable. Fortunately, on today's fast CPUs reduced frequency squares of order $n \leq 8$ can be enumerated directly. A simple backtracking program was used to compute the numbers in Table 2. The program was thoroughly checked for errors and most of the computations were performed at least twice. Therefore, we are confident all of the numbers in Table 2 are correct. For $n \leq 6$ our results agree with those of Brant and Mullen [1].

The results of our computations can be checked against each other by computing the total number of frequency squares in two ways. By [1, Theorem 2] we have:

n	λ	$\lvert f(n;\lambda)\rvert$	n	λ	$\lvert f(n;\lambda)\rvert$
7	$(5,2)$	9 876	8	$(5,1,1,1)$	40 171 008
	$(5,1,1)$	7 416		$(4,4)$	47 740 325
	$(4,3)$	98 484		$(4,3,1)$	771 067 692
	$(4,2,1)$	285 948		$(4,2,2)$	3 971 210 355
	$(4,1,1,1)$	214 752		$(4,2,1,1)$	3 166 707 276
	$(3,3,1)$	1 185 336		$(4,1,1,1,1)$	2 525 457 024
	$(3,2,2)$	4 582 740		$(3,3,2)$	20 826 177 696
	$(3,2,1,1)$	3 442 464		$(3,3,1,1)$	16 608 228 480
	$(3,1,1,1,1)$	2 586 432		$(3,2,2,1)$	85 538 838 240
	$(2,2,2,1)$	19 969 380		$(3,2,1,1,1)$	68 220 465 792
	$(2,2,1,1,1)$	14 998 608		$(3,1,1,1,1,1)$	54 413 316 096
	$(2,1,1,1,1,1)$	11 270 400		$(2,2,2,2)$	660 892 740 516
	$(1,1,1,1,1,1,1)$	16 942 080		$(2,2,2,1,1)$	527 062 142 160
8	$(6,2)$	318 930		$(2,2,1,1,1,1)$	420 413 623 488
	$(6,1,1)$	254 280		$(2,1,1,1,1,1,1)$	335 390 189 568
	$(5,3)$	12 268 464		$(1,1,1,1,1,1,1,1)$	535 281 401 856
	$(5,2,1)$	50 377 968			

Table 2. Number of reduced frequency squares.

$$|F(n;\lambda_1,\ldots,\lambda_s)| = \binom{n}{\lambda_1,\ldots,\lambda_s}\binom{n-1}{\lambda_1-1,\lambda_2,\ldots,\lambda_s}|f(n;\lambda_1,\ldots,\lambda_s)|.$$

On the other hand, this number can be computed from the complete list of isotopism class representatives:

$$|F(n;\lambda_1,\ldots,\lambda_s)| = |G_n|\cdot\sum_A \frac{1}{|\operatorname{Aut}(A)|}.$$

The sum is taken over all isotopism class representatives and $|\operatorname{Aut}(A)|$ denotes the autotopism group size, i.e. the number of isotopisms mapping the square A onto itself. If there are k distinct numbers among $\lambda_1,\ldots,\lambda_s$, occurring with frequencies $s_1,\ldots,s_k$, then $|G_n| = |S_n\times S_n\times(S_s)_\lambda| = (n!)^2 s_1!\cdots s_k!$. Autotopism group sizes were computed by B.D. McKay's `nauty` [5]. In each case the same total number of frequency squares was obtained, indicating that the results of our computations are consistent with each other.

4. The number of frequency squares of a given order

Brant and Mullen [1] noted that the number of frequency squares of order $n\le 6$ is a non-increasing function of the frequency vector. More precisely, frequency vectors are assumed to have non-increasing components and are compared lexicographically. In Theorem 3 of [1] it is stated that $|F(n;\lambda)|\ge|F(n;\lambda')|$ whenever $\lambda\le\lambda'$. While this still holds for $n=7$, it is no longer true for $n=8$. From the data in Table 2 we can compute $|F(8;4,4)| = 116963796250$; this number is smaller than both $|F(8;5,2,1)| = 888667355520$ and $|F(8;5,1,1,1)| = 2834466324480$. Another counterexample is $|F(8;3,3,2)| = 2449158497049600 < 3563924952268800 = |F(8;4,1,1,1,1)|$.

The main idea in the proof of Theorem 3 is to define a surjective mapping $\phi: F(n;\lambda)\to F(n;\lambda')$. The following concept is used, although it is not explicitly defined.

Definition 4.1. *Let $A\in F(n;\lambda_1,\ldots,\lambda_s)$ be a frequency square and $k\in\{1,\ldots,s\}$. A* k-transversal *of A is a set of n cells, one in every row and column, each containing the integer k.*

Obviously, Latin squares have a unique k-transversal for every k. Frequency squares also possess a k-transversal for every k, but it is not necessarily unique.

Lemma 4.2. *Let $A\in F(n;\lambda_1,\ldots,\lambda_s)$ be a frequency square. For every $k\in\{1,\ldots,s\}$ there is at least one k-transversal in A.*

Proof. Suppose $A = [a_{ij}]$; we define a bipartite graph with vertices $R_1, \ldots, R_n$ and $C_1, \ldots, C_n$. The pair (R_i, C_j) is joined by an edge if $a_{ij} = k$. Since k appears λ_k times in every row and column, all vertices have degree λ_k. By the marriage theorem the graph possesses a maximal matching, giving a k-transversal in A. □

In [1], the image $\phi(A)$ is obtained by repeatedly choosing a k-transversal and replacing k by some other integer in the corresponding cells. However, the definition of ϕ is flawed; the k-transversals are chosen by a greedy algorithm and this is not always possible without backtracking. Of course, one can use some other method to find k-transversals (e.g. the Ford-Fulkerson algorithm), but in general it is not possible to prove that ϕ is onto. Indeed, our counterexamples show that $\lambda \le \lambda'$ is not sufficient for a surjection $\phi : F(n;\lambda) \to F(n;\lambda')$ to exist.

The method of Brant and Mullen does give some information on the number of frequency squares. As a corollary of Theorem 3, it is stated that there are more Latin squares than any other kind of frequency squares of the same order. This can be proved by the ideas of [1].

Proposition 4.3. $|F(n;1,\ldots,1)| \ge |F(n;\lambda)|$ *for every frequency vector* λ.

Proof. We will define a surjective mapping $\phi : F(n;1,\ldots,1) \to F(n;\lambda)$. Let $\lambda = (\lambda_1, \ldots, \lambda_s)$ and define $\varphi : \{1,\ldots,n\} \to \{1,\ldots,s\}$ by

$$\varphi(i) = \begin{cases} i, & \text{for } i \le s \\ \min\{j \mid \lambda_1 + \ldots + \lambda_j - j \ge i - s\}, & \text{for } i > s. \end{cases}$$

The function ϕ acts on a Latin square $A = [a_{ij}]$ by substitution of the entries: $\phi(A) = [\varphi(a_{ij})]$. It remains to be shown that ϕ is onto.

Any frequency square $B \in F(n;\lambda)$ can be turned into a Latin square in the following way. Let $k = \min\{i \mid \lambda_i > 1\}$; choose a k-transversal in B and substitute $s+1$ instead of k in the corresponding cells. Thus, we get a frequency square in $F(n;1,\ldots,1,\lambda_k - 1, \lambda_{k+1}, \ldots, \lambda_s, 1)$. By repeated application of this transformation a Latin square A with the property $\phi(A) = B$ is obtained. □

References

[1] L.J. Brant and G.L. Mullen, Some results on enumeration and isotopic classification of frequency squares, Util. Math. **29** (1986), 231–244.

[2] C.J. Colbourn and J.H. Dinitz (eds.), The CRC handbook of combinatorial designs, CRC Press, Boca Raton, 1996.

[3] J. Dénes and A.D. Keedwell, Latin squares and their applications, Akadémiai Kiadó, Budapest, 1974.

[4] I.A. Faradžev, Constructive enumeration of combinatorial objects, Problèmes combinatoires et théorie des graphes, Colloq. Internat. CNRS **260**, Paris, 1978, pp. 131-135.

[5] B.D. McKay, `nauty` user's guide (version 1.5), Technical Report TR-CS-90-02, Department of Computer Science, Australian National University, 1990.

[6] B.D. McKay, Isomorph-free exhaustive generation, J. Algorithms **26** (1998), 306–324.

[7] B.D. McKay, A. Meynert, W. Myrwold, Small Latin squares, quasigroups and loops, preprint.

[8] B.D. McKay, I.M. Wanless, The number of Latin squares of order eleven, preprint.

[9] J.R. Nechvatal, Asymptotic enumeration of generalized Latin rectangles, Util. Math. **20** (1980), 273–292.

[10] R.C. Read, Every one a winner or how to avoid isomorphism search when cataloguing combinatorial configurations, Ann. Discrete Math. **2** (1978), 107-120.

[11] J. Shao, W. Wei, A formula for the number of Latin squares, Discrete Math. **110** (1992), 293–296.

Department of Mathematics
University of Zagreb
Bijenička 30
HR-10002 Zagreb, Croatia

E-mail address: `krcko@math.hr`