Sudoku, Mathematics and Statistics

Peter J. Cameron
Queen Mary University of London
p.j.cameron@qmul.ac.uk

Prospects in Mathematics, December 2006

Sudoku

> There's no mathematics involved. Use logic and reasoning to solve the puzzle.
>
> Instructions in *The Independent*

But who invented Sudoku?

- Leonhard Euler
- W. U. Behrens
- John Nelder
- Howard Garns
- Robert Connelly

Euler

Euler posed the following question in 1782.

> Of 36 officers, one holds each combination of six ranks and six regiments. Can they be arranged in a 6×6 square on a parade ground, so that each rank and each regiment is represented once in each row and once in each column?

NO!!

Latin squares

A *Latin square* of order n is an $n \times n$ array containing the symbols $1, \ldots, n$ such that each symbol occurs once in each row and once in each column.

The Cayley table of a group is a Latin square. In fact, the Cayley table of a binary system $(A, \circ)$ is a Latin square if and only if $(A, \circ)$ is a *quasigroup*. (This means that left and right division are uniquely defined, i.e. the equations $a \circ x = b$ and $y \circ a = b$ have unique solutions x and y for any a and b.)

Example 1.

$\circ$	a	b	c
a	b	a	c
b	a	c	b
c	c	b	a

About Latin squares

There is still a lot that we don't know about Latin squares.

- The number of different Latin squares of order n is not far short of n^{n^2} (but we don't know exactly). (By contrast, the number of groups of order n is at most about $n^{c(\log_2 n)^2}$, with $c = \frac{2}{27}$.)
- There is a Markov chain method to choose a random Latin square. But we don't know much about what a random Latin square looks like.
- For example, the second row is a permutation of the first; this permutation is a *derangement*

(i.e. has no fixed points). Are all derangements roughly equally likely?

Orthogonal Latin squares

Two Latin squares A and B are *orthogonal* if, given any k, l, there are unique i, j such that $A_{ij} = k$ and $B_{ij} = l$.

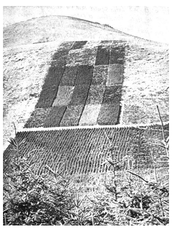
Euler was right that there do not exist orthogonal Latin squares of order 6; they exist for all other orders greater than 2.

But we don't know

- how many orthogonal pairs of Latin squares of order n there are;
- the maximum number of mutually orthogonal Latin squares of order n;
- how to choose at random an orthogonal pair.

Latin squares in statistics

Latin squares are used to "balance" treatments against systematic variations across the experimental layout.



A Latin square in Beddgelert Forest, designed by R. A. Fisher.

Behrens

The German statistician W. U. Behrens invented *gerechte designs* in 1956.

Take an $n \times n$ grid divided into n regions, with n cells in each. A gerechte design for this partition involves filling the cells with the numbers $1, \ldots, n$ in such a way that each row, column, or region contains each of the numbers just once. So it is a special kind of Latin square.

Example 2. Suppose that there is a boggy patch in the middle of the field.

1	2	3	4	5
4	5	1	2	3
2	3	4	5	1
5	1	2	3	4
3	4	5	1	2

Nelder

The statistician John Nelder defined a *critical set* in a Latin square in 1977. This is a partial Latin square which can be completed in only one way.

A *trade* in a Latin square is a collection of entries which can be "traded" for different entries so that another Latin square is formed.

A subset of the entries of a Latin square is a critical set if and only if it intersects every trade.

What is the size of the smallest critical set in an $n \times n$ Latin square? It is conjectured that the answer is $\lfloor n^2/4 \rfloor$, but this is known only for $n \leq 8$.

How difficult is it to recognise a critical set, or to complete one?

Garns

It was Howard Garns, a retired architect, who put the ideas of Nelder and Behrens together and turned it into a puzzle in 1979, in *Dell Magazines*.

A Sudoku puzzle is a critical set for a gerechte design for the 9×9 grid partitioned into 3×3 subsquares. The puzzler's job is to complete the square.

Garns called his puzzle "number place". It became popular in Japan under the name "Sudoku" in 1986 and returned to the West a couple of years ago.

Connelly

Robert Connelly proposed a variant which he called *symmetric Sudoku*. The solution must be a gerechte design for all these regions:

3	5	9	2	4	8	1	6	7
4	8	1	6	7	3	5	9	2
7	2	6	9	1	5	8	3	4
8	1	4	7	3	6	9	2	5
2	6	7	1	5	9	3	4	8
5	9	3	4	8	2	6	7	1
6	7	2	5	9	1	4	8	3
9	3	5	8	2	4	7	1	6
1	4	8	3	6	7	2	5	9

Rows Columns Subsquares
Broken rows Broken columns Locations

Coordinates

We coordinatise the cells of the grid with F^4, where F is the integers mod 3, as follows:

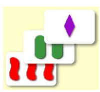
- the first coordinate labels large rows;
- the second coordinate labels small rows within large rows;
- the third coordinate labels large columns;
- the fourth coordinate labels small columns within large columns.

Now Connelly's regions are cosets of the following subspaces:

Rows	$x_1 = x_2 = 0$	Columns	$x_3 = x_4 = 0$
Subsquares	$x_1 = x_3 = 0$	Broken rows	$x_2 = x_3 = 0$
Broken columns	$x_1 = x_4 = 0$	Locations	$x_2 = x_4 = 0$

Affine spaces and SET®

The card game SET has 81 cards, each of which has four attributes taking three possible values (number of symbols, shape, colour, and shading). A winning combination is a set of three cards on which either the attributes are all the same, or they are all different.



Each card has four coordinates taken from F (the integers mod 3), so the set of cards is identified with the 4-dimensional affine space. Then *the winning combinations are precisely the affine lines!*

Perfect codes

A *code* is a set C of "words" or n-tuples over a fixed alphabet F. The *Hamming distance* between two words v, w is the number of coordinates where they differ; that is, the number of errors needed to change the transmitted word v into the received word w.

A code C is *e-error-correcting* if there is *at most* one word at distance e or less from any codeword. [Equivalently, any two codewords have distance at least $2e + 1$.] We say that C is *perfect e-error-correcting* if "at most" is replaced here by "exactly".

Perfect codes and symmetric Sudoku

Take a solution to a symmetric Sudoku puzzle, and look at the set S of positions of a particular symbol s. The coordinates of the points of S have the property that any two differ in at least three places; that is, they have Hamming distance at least 3. [For, if two of these words agreed in the positions 1 and 2, then s would occur twice in a row; and similarly for the other pairs.]

Counting now shows that any element of F^4 lies at Hamming distance 1 or less from a unique element of S; so S is a perfect 1-error-correcting code.

So a symmetric Sudoku solution is a partition of F^4 into nine perfect codes.

All symmetric Sudoku solutions

Now it can be shown that a perfect code C in F^4 is an *affine plane*, that is, a coset of a 2-dimensional subspace of F^4. To show this, we use the *SET® principle*: We show that if $v, w \in C$, then the word which agrees with v and w in the positions where they agree and differs from them in the positions where they differ is again in C.

So we have to partition F^4 into nine special affine planes.

It is not hard to show that there are just two ways to do this.

One solution consists of nine cosets of a fixed subspace.

There is just one further type, consisting of six cosets of one subspace and three of another. [Take a solution of the first type, and replace three affine planes in a 3-space with a different set of three affine planes.]

All Sudoku solutions

By contrast, Jarvis and Russell showed that the number of different types of solution to ordinary Sudoku is 5 472 730 538.

They used the *Orbit-Counting Lemma*:

> the number of orbits of a group on a finite set is equal to the average number of fixed points of the group elements.

An earlier computation by Felgenhauer and Jarvis gives the total number of solutions to be 6 670 903 752 021 072 936 960. Now for each conjugacy class of non-trivial symmetries of the grid, it is somewhat easier to calculate the number of fixed solutions.

Some open problems

Given a $n \times n$ grid partitioned into n regions each of size n:

- What is the computational complexity of deciding whether there exists a gerechte design?
- Assuming that there exists a gerechte design, how many are there (exactly or asymptotically), and how do we choose one uniformly at random?
- Assuming that there exists a gerechte design, what is the maximum number of pairwise orthogonal gerechte designs?
- Which gerechte designs have "good" statistical properties?

If we are given a Latin square L, and we take the regions to be the positions of symbols in L, then a gerechte design is a Latin square orthogonal to L; so the above questions all generalise classical problems about orthogonal Latin squares.

The last two questions are particularly interesting in the case where $n = kl$ and the regions are $k \times l$ rectangles.

References

- R. A. Bailey, P. J. Cameron and R. Connelly, Sudoku, Sudoku, gerechte designs, resolutions, affine space, spreads, reguli, and Hamming codes, *American Math. Monthly*, to appear. Preprint available from `http://www.maths.qmul.ac.uk/~pjc/preprints/sudoku.pdf`