association scheme

Canonical name	AssociationScheme
Date of creation	2013-03-22 19:15:03
Last modified on	2013-03-22 19:15:03
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	10
Author	CWoo (3771)
Entry type	Definition
Classification	msc 05E30
Defines	associate class
Defines	rank
Defines	intersection number
Defines	valency

Association schemes were introduced by statisticians in the 1950's to analyze designs of statistical experiments. Today, association schemes are useful not only in experimental designs, but in other areas of mathematics such as combinatorics (coding theory) and group theory (permutation groups).

There are several equivalent ways to define an association schemes. Three useful ones are illustrated here:

Let Ω be a non-empty set with n elements, and s a positive integer.

Definition 1. An *association scheme* $\mathscr{Q}$ on Ω is a partition on $\Omega \times \Omega$ into sets $C_0, C_1, \ldots, C_s$ called *associate classes*, such that

- each C_i is a symmetric relation on Ω, and C_0 in particular is the diagonal relation,
- for $i, j, k \in \{0, 1, \ldots, s\}$, there is an integer p_{ij}^k such that, for any $(a, b) \in C_k$,
$$|\{c \in \Omega \mid (a,c) \in C_i \text{ and } (c,b) \in C_j\}| = p_{ij}^k.$$

If we write $C(a, b; i, j)$ for the set $\{c \in \Omega \mid (a,c) \in C_i \text{ and } (c,b) \in C_j\}$, then the second condition says that for any $(a, b) \in C_k$, the value $|C(a, b; i, j)|$ is a constant, depending only on i, j and k, and not on the particular elements of C_k. This implies that, for any $i, j \in \{0, 1, \ldots, n\}$, the relation $C_i \circ C_j$ is a union of (some of) the C_k's.

The definition above can be restated in graph theoretic terminology. First, think of Ω is a set of vertices, and two-element subsets of Ω are edges. The complete graph on Ω is just the set of all two-element subsets of Ω. We may color the edges of the graph. Say there are colors labeled 1 through s. For each color i, let C_i be the set of edges with color i. Then each C_i is just a symmetric relation on Ω, and that all the C_i's, together with the diagonal relation, partition the set $\Omega \times \Omega$. This is basically the first condition of the definition above. In this regard, we can redefine an association scheme graph theoretically, as follows:

Definition 2. An *association scheme* $\mathscr{Q}$ is a surjective coloring on the edge set of a complete graph whose vertex set is Ω, by a set of s colors (numbered 1 through s), such that

> for any $i, j, k \in \{1, \ldots, s\}$, there is an integer p_{ij}^k such that if $\mathscr{Q}(a, b) = k$ (the edge $\{a, b\}$ has color k), then
> $$|\{c \in \Omega \mid \mathscr{Q}(a,c) = i \text{ and } \mathscr{Q}(c,b) = j\}| = p_{ij}^k.$$

In words, the definition says that, for any color k, and any given edge e with color k, the number of triangles (a triangle in a graph is a cycle consisting of three edges) with e as an edge, and two other edges with colors i, j respectively, is p_{ij}^k.

The first definition can also be viewed in terms of matrices, and adjacency matrices more specifically. Given a finite set Ω, a binary relation R on Ω naturally corresponds to matrix A called the adjacency matrix of R. Entry (i, j) is 1 if the i-th element and the j-th element are related by R, and 0 otherwise. If R is reflexive, then A has all 1's in its diagonal, and if R is symmetric, then so is A. Also, it is easy to see that the composition of two binary relations is the same as the product of their corresponding adjacency matrices. Then the comment in the paragraph after the first definition is the same as saying that the adjacency matrix of $C_i \circ C_j$ is a linear combination of the adjacency matrices of $C_0, C_1, \ldots, C_s$. This gives us the third definition below:

Definition 3. An *association scheme* is a finite set $\mathscr{Q}$ of $n \times n$ non-zero matrices $A_0, A_1, \ldots, A_s$ whose entries are 0's and 1's, such that

- each A_i is a symmetric matrix, with $A_0 = I_n$, the identity matrix,
- $A_0 + A_1 + \cdots + A_s = J_n$, the matrix whose entries are all 1's, and
- for any $i, j \in \{0, \ldots, s\}$, $A_i A_j$ is a linear combination of $A_0, A_1, \ldots, A_s$.

By the definitions of the matrices A_i and the second condition, for every pair (r, s), exactly one of the $s+1$ matrices has 1 in cell (r, s), and all others have 0 in the corresponding cell. As a result, the $s + 1$ matrices are linearly independent.

Also, in view of the discussion above, it is easy to see that

$$A_i A_j = p_{ij}^0 A_0 + p_{ij}^1 A_1 + \cdots + p_{ij}^s A_s.$$

Some terminology. s is called the *rank* of the association scheme $\mathscr{Q}$. Any $a \in \Omega$, an element $c \in \mathscr{Q}$ is said to be an *i-associate* of a if $(a, c) \in C_i$. For each $a \in \Omega$, define:

$$C_i(a) := \{c \in \Omega \mid (a, c) \in C_i\}.$$

So $C_i(a)$ is the set of all i-associates of a. Then

$$|C_i(a) \cap C_j(b)| = p_{ij}^k, \text{ whenever } (a, b) \in C_k,$$

and because of the above equation, each p_{ij}^k is called an *intersection number* of $\mathscr{Q}$. For each i, the intersection number p_{ii}^0 is called the *valency* of C_i, denoted by a_i.

Some basic properties of the intersection numbers:

- $p_{ij}^k = p_{ji}^k$
- $$p_{i0}^k = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases}$$
- $|C_i(c)| = p_{ii}^0 = a_i$ for all $c \in \Omega$.

References

[1] R. A. Bailey, *Association Schemes, Designed Experiments, Algebra and Combinatorics*, Cambridge University Press (2004)