ON THE LINEAR INDEPENDENCE OF SETS OF $2^q$ COLUMNS OF CERTAIN $(1, -1)$ MATRICES 
WITH A GROUP STRUCTURE, AND ITS CONNECTION WITH FINITE GEOMETRIES

J. N. Srivastava

Department of Statistics
Colorado State University, Fort Collins, CO 80521 U.S.A.

ABSTRACT

Consider a set of $m$ symbols (indeterminates) $F_1, \ldots, F_m$, and let $G$ be the group of order $2^m$ generated by multiplying these symbols two, or three, or more at a time, where the multiplication is assumed commutative, and where $F_j^2 = \mu$ (the identity element of $G$) for all $j$. The elements of $G$ can be written, in order, as $(\mu; F_1, \ldots, F_m; F_1F_2, F_1F_3, \ldots, F_{m-1}F_m; F_1F_2F_3, \ldots; F_1F_2 \ldots F_m)$. Consider a matrix $A(N \times 2^m)$ over the real field whose columns correspond in order to the elements of the group $G$. The elements of $A$ are $1$ and $(-1)$, and are obtained as follows. The elements of $A$ in the column corresponding to $\mu$ are all equal to $1$. The next $m$ columns of $A$, filled in arbitrarily, constitute an $(N \times m)$ submatrix, say $A^\star$. Finally, for all $f$ ($1 < f < m$), and all $i_1, \ldots, i_f$ (with $1 < i_1 < i_2 < \ldots < i_f < m$), the column of $A$ corresponding to $F_{i_1}F_{i_2} \ldots F_{i_f}$ is obtained by taking the Schur product of the columns of $A$ (or $A^\star$) corresponding to $F_{i_1}, F_{i_2}, \ldots, F_{i_f}$. The matrix $A$ (over the real field) is said to have the property $P_t$ if and only if every set of $t$ columns of $A$ is linearly independent. In this paper, for all positive integers $q$, we obtain necessary conditions on $A^\star$ such that every $(N \times 2^q)$ submatrix $A^{**}$ in $A$ has rank $2^q$. A non-statistical introduction together with an illustrative example is provided.

INTRODUCTION

We first execute the remark made in the last sentence above.

This subject is a part of the theory of "the design of factorial experiments of the $2^m$ type." Here, we are concerned with (statistically) planning a scientific experiment in which we are studying the effect of $m$ factors (or variables) each at two levels, on some characteristic of the experimental material. For example, we may have an agricultural experiment with $4$ ($=m$) factors, these being the nitrogen, phosphorus, potassium, and organic fertilizers, and the characteristic under study may be the yield of wheat. The two levels of each fertilizer (indicated by $1$ and $(-1)$ respectively) may indicate the presence and absence respectively of the fertilizer. Each row of $A^\star$ then indicates a particular treatment-combination, i.e. a combination of levels of these factors. The elements of $G$ can be interpreted as the names of certain parameters describing the effect of the various fertilizers on the yield of wheat. Thus, $\mu$ denotes the over-all average of the effects of the various treatment-combinations, $F_i$ the main effect of the $i$th factor, $F_iF_j$, the interaction between
the ith and jth factors, $F_{ij}F_{jk}$, the three-factor interaction between the ith, jth, and kth factors, and so on. The effect of any particular treatment-combination (which corresponds to a row of $A^*$) is a linear function of the above parameters, the coefficients being the corresponding elements in the row of $A$ containing this particular row of $A^*$.

For any positive integer $t$, the significance of $A$ having the property $P_{2t}$ is as follows. Suppose no random fluctuations are present, and at most $t$ out of the $2^m$ parameters are non-zero. Also, assume that an experiment is done using the $N$ treatment-combinations represented by the $N$ rows of $A^*$. Then a necessary and sufficient condition that the value of the non-zero parameters can be determined precisely is that $A$ have property $P_{2t}$. Thus, this problem has a fundamental importance in the design of experiments, and is deeply connected with information and coding theory.

**Definition 1.1** (a) Let $T$ $(N \times m)$ be the $(0,1)$ matrix obtained from $A^*$ by replacing $(-1)$ by $0$. Then $T$ is called the design. (b) Let $\overline{T}$ be the matrix obtained from $T$ by interchanging 0 and 1; we shall consider $\overline{T}$ over $GF(2)$.

To help in clarifying ideas, we now present an example of the matrices $T$, $A$, etc. Thus, the matrix $T$ at (1.1) below represents a design for a $2^4$ factorial experiment, the rows of $T$ representing combinations of levels of the four factors. The matrix $A$ corresponding to $T$ is presented at (1.2). For convenience, the elements of the group $G$ corresponding to each column of $A$ is indicated at the top of the column:

\[
T = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 \\
0 & 1 & 1 & 0
\end{bmatrix}, \quad \overline{T} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1
\end{bmatrix}, \quad A^* = \begin{bmatrix}
+ & + & + & + \\
+ & - & - & - \\
+ & + & - & - \\
+ & - & + & + \\
+ & + & + & - \\
+ & - & - & + \\
+ & + & - & - \\
+ & - & + & +
\end{bmatrix}
\]

\[
A = \begin{bmatrix}
\mu & F_1 & F_2 & F_3 & F_4 & F_{12} & F_{13} & F_{14} & F_{23} & F_{24} & F_{34} & F_{123} & F_{124} & F_{134} & F_{234} & F_{1234}
\end{bmatrix}
\]

where, in the above, $(+)$ and $(-)$ stand respectively for $(+1)$ and $(-1)$. Notice that