PROPERTIES OF THE ITEM SCORE MATRIX

ANGUS G. MACLEAN
CALIFORNIA TEST BUREAU

A method of deriving from the item score matrix all the usual statistics describing the performance on a test of a group of examinees is given. Since this matrix usually is not actually written out, but is implicit in a set of punched cards, a method of working from a more compact matrix $F$ is described. A numerical example is presented. Applications and advantages of the method are cited, as compared with that of recording only the examinees' test scores and the item difficulties.

Equally Weighted Items

An item score matrix $(X)$ is an $N$ by $n$ rectangular matrix with elements $X_{si}$ all of which are either 1 or 0. Each row of $(X)$ is a row vector $(X_s)$, which lists the item scores of student $s$. If items are to be weighted equally the sum of the elements of $(X_s)$ is $\sum_{i=1}^{n} X_{si} = X_s$, the test score of student $s$. The sum of the test scores of all students in the sample is

$$\sum_{s=1}^{N} X_s = \sum_{s=1}^{N} \sum_{i=1}^{n} X_{si} = T,$$

the sum of all elements of $(X)$.

The column sums of $(X)$ are of interest since

$$\sum_{s=1}^{N} X_{si} = f_i,$$

the number of students responding correctly to item $i$.

The square of the test score for student $s$ is obtainable by premultiplying the row vector $(X_s)$ by its transpose, a procedure which yields a square symmetric matrix of unit rank:

$$(X_s^2) = (X_s)'(X_s).$$

The sum of all elements of this matrix is $X_s^2$.

Some of the operations to be discussed lead to scalar values, others to matrices, the sums of whose elements are those values. For the purposes of clarity, therefore, all symbols for matrices are enclosed in parentheses, while symbols not so enclosed will denote numbers.

The elements of $(X_s)'(X_s)$ are the products $X_{si}X_{si}$ for student $s$. Therefore

$$X_s^2 = \sum_{i} \sum_{i} X_{si}X_{si}.$$
In general, the square of a sum may be obtained by squaring the row vector whose elements are the sum's components, then summing the elements of the square matrix so obtained.

Summing (4) over the \( N \) students gives

\[
\sum_{i=1}^{N} X_i^2 = \sum_{i} \sum_{j} X_{ij}X_{ij} = S.
\]

\( S \) is also obtained by summing the elements of a square symmetric matrix \((S)\) obtained by

\[
(S) = (X)'(X).
\]

It could also be obtained by adding the \( N \) matrices \((X_i^2)\) obtained by (3), that is,

\[
(S) = (X)'(X) = \sum_{i=1}^{N} (X_i)'(X_i).
\]

The side elements of \((S)\) are the cross-product sums \( S_{ij} \) of the columns of \((X)\), while the diagonal elements \( S_i \) are the result of multiplying the columns by themselves. That is,

\[
S_i = \sum_j X_{ij}^2,
\]

\[
S_{ij} = \sum_j X_{ij}X_{ji}.
\]

\( T \) and \( S \) always denote summation over the \( N \) individuals in the sample. They are the statistics used in calculating standard deviations and correlations, as follows:

\[
\sigma_i = \frac{\sqrt{L_i}}{N},
\]

\[
r_{ij} = \frac{L_{ij}}{\sqrt{L_i} \sqrt{L_j}},
\]

in which

\[
L_i = NS_i - T_i^2,
\]

\[
L_{ij} = NS_{ij} - T_i T_j.
\]

It so happens, when scores are either 1 or 0, that

\[
S_i = T_i = f_i,
\]

and

\[
S_{ij} = f_{ij}.
\]