PSYCHOMETRIKA—VOL. 54, NO. 2, 231–235
JUNE 1989

CONVERGENCE PROPERTIES OF AN ITERATIVE PROCEDURE OF IPSATIZING AND STANDARDIZING A DATA MATRIX, WITH APPLICATIONS TO PARAFAC/CANDECOMP PREPROCESSING

JOS M. F. TEN BERGE AND HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

Centering a matrix row-wise and rescaling it column-wise to a unit sum of squares requires an iterative procedure. It is shown that this procedure converges to a stable solution. This solution need not be centered row-wise if the limiting point of the iterations is a matrix of rank one. The results of the present paper bear directly on several types of preprocessing methods in Parafac/Candecomp.

Key words: ipsative data, doubly standardized matrices.

Clemans (1966) has given an extensive discussion of ipsative matrices, defined as matrices having the property that the sum of the scores over the attributes for each of the entities is a constant. Typically, ipsative matrices are obtained by row-wise centering, that is, subtracting the means row-wise, where each row corresponds to an entity. In addition, Clemans has suggested the possibility of standardizing the ipsative matrix column-wise, which yields the "ipsative-standard score matrix" (p. 7). However, it is readily seen that the latter matrix is no longer ipsative, because rescaling a matrix column-wise affects the row means differentially. Therefore, if a matrix is desired which is both ipsative and standardized, then one has to resort to an iterative procedure (Cattell, 1966, p. 118). Procedures of this kind have recently drawn attention in the context of Parafac/Candecomp (Harshman & Lundy, 1984; Kruskal, 1984). In particular, iteratively centering a three-mode array of order $p \times q \times r$ *across* the p elements of one mode and rescaling *within* elements of that mode comes down to the iterative process of centering a matrix of order $p \times qr$ column-wise and rescaling it row-wise. Kruskal has recently claimed that "though it has not been proved mathematically, we believe that this procedure will always converge" (p. 61). In the present paper, a mathematical proof will be provided in support of this claim. However, it will also be shown that, after convergence, the resulting matrix need not be ipsative and standardized at the same time.

Formal Statement of the Problem, and a Basic Result

Let Y denote a $k \times n$ data matrix with observations for k entities on n variables and let $J = (I - 11'/n)$, the centering operator of order n. Then YJ is the row-wise centered version of Y. Rescaling the columns of Y to unit sums of squares can be expressed as post-multiplication of YJ by $D^{-1/2}$, where D is the diagonal of $(JY'YJ)$. The resulting matrix $YJD^{-1/2}$ can again be centered row-wise by taking $YJD^{-1/2}J$, and so on.

It is important to note that this procedure consists of subsequent postmultiplications of Y. Therefore, if the initial Y is centered *column-wise*, this property will survive each additional step of the procedure. For this reason, rescaling our matrix column-

Requests for reprints should be sent to Jos M. F. ten Berge, Department of Psychology, University of Groningen, Grote Markt 32, 9712 HV Groningen, THE NETHERLANDS.

0033-3123/89/0600-9059 $00.75/0

wise is tantamount to standardizing it column-wise, provided that the initial Y has been centered column-wise.

The main purpose of the present paper is to show that the iterative procedure converges. This will be accomplished by showing that the procedure monotonely decreases the nonnegative function $1'Y_j'Y_j1$, evaluated after the rescaling step, where Y_j is the current version of Y after j iterations. First, however, we need to have a general result on correlation matrices.

Result 1. Let R be an $n \times n$ correlation matrix and let R_* be the correlation matrix obtained from JRJ by

$$R_* = D^{-1/2}JRJD^{-1/2}, \tag{1}$$

where $J = J^2$ is the $n \times n$ centering matrix and $D = \text{Diag}(JRJ)$. Then

$$1'R1 \geq 1'R_*1. \tag{2}$$

Proof. Let Y be any matrix such that $Y'Y = R$, and let e_i be the i-th column of the $n \times n$ identity matrix. From the Schwartz inequality we have, for $i = 1, \ldots, n$,

$$(e_i'JRe_i)^2 = \{(e_i'JY')(Ye_i)\}^2 \leq (e_i'JRJe_i)(e_i'Re_i) = e_i'JRJe_i, \tag{3}$$

and hence

$$(JR)_{ii} \leq D_{ii}^{1/2}. \tag{4}$$

Noting that $D^{1/2} = \text{Diag}(JRJD^{-1/2})$ it can be seen from (4) that

$$\text{tr } JRJD^{-1/2} \geq \text{tr } JR. \tag{5}$$

Furthermore, we have the inequality

$$\|YJ - YJD^{-1/2}J\|^2 \geq 0, \tag{6}$$

which is equivalent to

$$\text{tr } JR + \text{tr } JR_* \geq 2 \text{ tr } JRJD^{-1/2}. \tag{7}$$

Combining (5) and (7) yields

$$\text{tr } JR_* \geq \text{tr } JR. \tag{8}$$

By expanding the definition of J we arrive at

$$1'R_*1 \leq 1'R1, \tag{9}$$

which completes the proof of Result 1. □

Result 1 is the key result to be used in the convergence proof below. However, it is also interesting in its own right. That is, centering a matrix of standard scores row-wise reduces the sum of the correlations between the variables, according to Result 1.

A Proof of Convergence

Let the iterative procedure, described above, be summarized by the expression

$$Y_{j+1} = Y_j J D_j^{-1/2}, \qquad j = 0, 1, 2, 3, \ldots \tag{10}$$

where D_j = [Diag $(JY_j'Y_jJ)$] and the initial matrix $Y = Y_0$ is understood to have unit sums of squares column-wise. In (10) and elsewhere in the present paper, it is tacitly assumed that D_j is nonsingular. This excludes the case where one or more columns of Y_jJ are zero.

Define, for $j \geq 0$, $R_j = Y_j'Y_j$ and $R_{j+1} = Y_{j+1}'Y_{j+1}$. Then it is clear from Result 1 that the iterative procedure (10) reduces $1'R_j1$ monotonely, as j increases. Because $1'R_j1$ is bounded below (by zero), the iterative procedure must converge to a solution where the difference between $1'R_j1$ and $1'R_{j+1}1$ tends to zero. It will now be shown that the difference between Y_j and Y_{j+1} tends to zero as well. For this purpose, let us revisit the proof of Result 1. Suppose that equality holds in (9). Then we have, with $Y_* = YJD^{-1/2}$,

$$YJ = Y_*J, \tag{11}$$

see (6). It follows that

$$D = \text{Diag}\,(JY'YJ) = \text{Diag}\,(JY_*'Y_*J) = D_*. \tag{12}$$

Rescaling the columns of Y_*J again to unit sums of squares results in the matrix

$$Y_*JD_*^{-1/2} = Y_*JD^{-1/2} = YJD^{-1/2} = Y_*. \tag{13}$$

This shows that the procedure (10) converges to a stable matrix Y_* with unit sum of squares for each column.

Unfortunately, this matrix Y_* need not be centered row-wise. Note that it has not been shown that $1'R1$ tends to zero, therefore $\|Y_*J - Y_*\|^2$ has not been shown to tend to zero. This means that a solution may be obtained which fails to meet the purpose of the procedure.

In fact, examples can be constructed where the final solution Y_* is stable, with Y_*J also stable, but $Y_* \neq Y_*J$. However, these examples have a peculiar form of necessity, as will be shown below.

Result 2. If $Y_j = Y_{j+1}$ and $Y_j \neq Y_jJ$ then Y_j has rank one.

Proof: If $Y_j = Y_{j+1}$ then $R_j = R_{j+1}$ and we have $1'R_j1 = 1'R_{j+1}1$. In the proof of Result 1 we cannot have equality in (9) unless (3) holds as an equality for $i = 1, \ldots, n$. In the present notation, this yields

$$Y_jJe_i = \lambda_i Y_je_i \tag{14}$$

for certain scalars λ_i, $i = 1, \ldots, n$, if $Y_j = Y_{j+1}$. Defining $\mu_i = n(1 - \lambda_i)$ we can simplify (14) as

$$\mu_i Y_j e_i = Y_j 1 \tag{15}$$

for $i = 1, \ldots, n$. If, for some i, $\mu_i = 0$, then $Y_j1 = 0$ and $Y_j = Y_jJ$, which would produce a contradiction. Therefore, $\mu_i \neq 0$ for $i = 1, \ldots, n$. This shows that all columns of Y_j are proportional so Y_j has rank 1. This completes the proof of Result 2. □

Conversely, when Y_j has rank one then $Y_j = Y_{j+1}$ in any case. This follows at once upon writing $Y_j = xu'$ for some vector x with $x'x = 1$ and a sign vector u. However, if the rank of Y_j is 1 and, therefore, $Y_j = Y_{j+1}$, we still need not have $Y_jJ = Y_j$. It follows that the procedure may converge to a rescaled yet noncentered solution. It is clear from Result 2 that this can only happen if the procedure converges to a rank one solution.

Result 3. The rank of Y_j $(j = 1, 2, \ldots)$ is at least (rank $Y_0 - 1$).

Proof. After j iterations we have

$$Y_j D_{j-1}^{1/2} = Y_0 J D_0^{-1/2} J D_1^{-1/2} \cdots D_{j-2}^{-1/2} J \equiv Y_0 W \tag{16}$$

for some $n \times n$ matrix W. It is well-known that J has $n - 1$ unit eigenvalues and one zero eigenvalue. Therefore, J can be written as $J = KK'$ for some $n \times (n - 1)$ matrix K with $K'K = I_{(n-1)}$. Let $K'D_j^{-1/2}K \equiv W_j$. Then $Y_0 W$ can be written as

$$Y_0 W = Y_0 K W_0 W_1 \cdots W_{j-2} K'. \tag{17}$$

Clearly, each W_j is nonsingular hence W is of rank $n - 1$. By Sylvester's inequality (Gantmacher, 1959, p. 66) we have

$$\text{rank } Y_0 W \geq (\text{rank } Y_0) + (\text{rank } W) - n = (\text{rank } Y_0) - 1. \tag{18}$$

Clearly, rank $Y_0 W$ = rank Y_j. This completes the proof of Result 3. □

From Result 2 and Result 3 one might be tempted to infer that starting with an initial Y_0 that has rank at least three will prevent convergence to a noncentered solution. However, a reviewer pointed out that this does not follow. Specifically, suppose that Y_0 has rank three or higher. Then the solution has at least rank two. However, the procedure is terminated as soon as $1'R_j 1 = 1'R_{j+1} 1$ holds in a finite number of decimal places, implying that (14) and (15) do not hold exactly. Therefore, the solution does not satisfy (14) and (15) yet the limiting point of the Y_j as j tends to infinity may satisfy (14) and (15). In other words, if the procedure converges to a solution of rank one the actual solution Y_j may have higher rank. In fact, the reviewer has provided examples where Y_0 has arbitrary rank yet the procedure yields a noncentered solution. It can be concluded that the iterative procedure (10) may yield a non-centered solution if it converges to a rank one matrix. The problem, how to determine from Y_0 whether or not the procedure will converge to a rank one solution, remains yet to be solved.

Discussion

The results of the present paper are particularly relevant for certain varieties of Parafac preprocessing, where the three-mode array is to be centered across elements of one or two modes, and to be rescaled *within* elements of either mode, see Kruskal (1984) and Harshman and Lundy (1984). However, many other types of Parafac preprocessing remain yet to be examined (Harshman & Lundy, p. 247–271).

Harshman and Lundy (1984, p. 252) have given a matrix expression for the iterative centering and rescaling of two modes of a three-way array, where each slab of the array is premultiplied by the same matrix M_a and postmultiplied by the same matrix M_b'. They note that the properties of these matrices "are still under study". The proof of Result 3 has an implication for these matrices. That is, they have precisely the same structure as matrix W in (17), hence they are of rank $k - 1$ and $n - 1$ respectively, if the slabs are of order $k \times n$. Combining this with the observation that both M_a and M_b are column-centered (Harshman & Lundy, p. 253) it can be concluded that the column spaces of M_a and M_b are known.

References

Cattell, R. B. (1966). The data box: Its ordering of total resources in terms of possible relational systems. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 67–128). Chicago: Rand McNally.

Clemans, W. V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometrika Monograph No. 14*, *31*(4, Pt. 2).

Gantmacher, F. R. (1959). *The theory of matrices (Vol. 1)*. New-York: Chelsea.

Harshman, R. A., & Lundy, M. E. (1984). Data preprocessing and the extended Parafac model. In H. G. Law, C. W. Snyder, J. A. Hattie, & R. P. Mc Donald (Eds.), *Research methods for multimode data analysis* (pp. 216–284). New-York: Praeger.

Kruskal, J. B. (1984). Multilinear methods. In H. G. Law, C. W. Snyder, J. A. Hattie, & R. P. Mc Donald (Eds.), *Research methods for multimode data analysis* (pp. 36–62). New-York: Praeger.

Manuscript received 1/7/87
Final version received 12/4/87