Converting an Inconsistent Matrix into a Consistent Non-Singular Matrix*

Marion F. Shaycoft
American Institutes for Research

A. The Inconsistent-Matrix Problem

A correlation matrix in which every correlation coefficient has been computed directly from the raw data and is based on exactly the same cases is internally consistent (in other words Gramian) automatically.

Unfortunately not all correlation matrices used in research can fit this neat pattern. Sometimes, for instance, one is up against a severe missing data problem and is forced to base each correlation coefficient on all cases having data available for that coefficient--whether or not they have valid data on all of the other variables included in the matrix. And sometimes even when there isn't a missing data problem, the initial correlation matrix, which is internally consistent, may have to be subjected subsequently to some sort of statistical manipulation that is desirable for other reasons but has the unfortunate effect of introducing slight inconsistencies into the results. One such kind of statistical manipulation is correction of correlation coefficients for attenuation. Although this won't <u>necessarily</u> cause a matrix to become inconsistent, it is quite likely to have this effect if the reliability coefficients that are used for the purpose are underestimates of the actual reliability of the variable. Another kind of statistical manipulation that may on occasion produce an inconsistent matrix occurs when one has two or more consistent matrices, each based on the same variables but on different cases, and these matrices are then combined, via Fisher's z, to obtain a single matrix representing the correlations for the "average group". The average matrix can be inconsistent.

Any correlation used for multiple regression, canonical correlation, factor analysis, discriminant functions, or any other form of multivariate analysis should be a consistent matrix, at least in its initial state with

*Paper presented at Annual Meeting of American Educational Research Association, in Chicago on 9 February 1968.

ones in the diagonal*. This is not merely desirable; in most situations it is virtually essential (at least in my opinion). An inconsistent matrix may result in multiple or partial correlations outside the +1 to -1 range to which correlation coefficients are restricted. And if it contains zero-order correlations greater than 1, (as may happen in the case of overcorrection for attenuation), it may result in multiple or partial correlations that are imaginary numbers. But even if no such blatant illogicalities as coefficients that are imaginary or greater than 1 manifest themselves to the naked eye, the fact that illogicalities are necessarily inherent in any inconsistent matrix renders multivariate analysis based on such a matrix suspect.

That leaves us with the practical problem that sometimes the imperatives of research force us to cope with an inconsistent matrix.

B. Solving the Problem

The best way to cope with an inconsistent matrix is to find a way of making minimal changes in it that will convert it to a consistent matrix usable in multivariate analysis.

A procedure that I developed** a while ago to do just this worked very well in general, but had one awkward feature; the corrected matrix was necessarily singular. This is a drawback because use of such a matrix is inconvenient, or in some situations impossible. If the matrix is singular, it is generally necessary to eliminate one or more variables, to make it non-singular, before multivariate analysis can be undertaken.

---

*When values other than ones--communality estimates in factor analysis for instance--are to be substituted in the diagonal, the statement about the importance of a strictly Gramian matrix is intended to apply only to the initial correlation matrix, with 1's in the diagonal, since it is recognized that use of communality estimates may make the matrix very slightly inconsistent.

**This procedure was described in an earlier paper, "Method of Making an Inconsistent Matrix Consistent," presented at the American Psychological Association Convention, in Washington, D.C. on 3 September 1967.

The purpose of this paper is to describe a modification of the earlier procedure which eliminates its drawback; the modified procedure produces a matrix that is almost certain to be non-singular.

C. The Procedure

1. Determining whether the matrix is inconsistent

Whether the matrix is consistent can be determined easily from its eigenvalues. A matrix that has any negative eigenvalues is necessarily inconsistent (non-Gramian); conversely if the matrix is inconsistent it will have at least one negative eignevalue.

2. The earlier procedure for correction

To understand the new procedure, which produces a corrected matrix that is non-singular, it is necessary to understand the earlier procedure, which produced a singular matrix. What it amounts to, basically, is just the reverse of factor analysis. In factor analysis one breaks a set of correlated variables down into a set of factor loadings which could explain the correlations. In the reverse procedure one takes a set of factor loadings and from it builds up the correlation matrix compatible with it. The procedure is summarized in Table 1. The rank of the final (consistent) matrix is equal to the number of positive eigenvalues for the original (inconsistent) matrix. In other words the number of positive eigenvalues remains the same, while the correction-for-inconsistency process causes each negative eigenvalue to be replaced by a zero eigenvalue.

Each zero eigenvalue represents a linear dependency--in other words a fixed linear relationship linking two or more variables and applying without exception. Before the consistent correlation matrix can be used in multivariate analysis it is usually necessary to eliminate enough variables to get rid of each fixed linear relationship. Fortunately there is considerable leeway in choosing what variables to eliminate, since

Table 1. Condensed description of original and revised procedures for making an inconsistent correlation matrix consistent

1. Starting point: An inconsistent correlation matrix	M_n = inconsistent n-variable matrix (with i's in diagonal), corresponding to which a consistent matrix is desired.	
2. Matrix to which correction for . . .	M = inconsistent n´-variable matrix which is identical with M_n except that some dummy variables may have been added.	
	$k = 0$ (i.e., no dummy variables) $n' = n$ $M = M_n$	$k \geq n' - m$
3. Partial matrix of factor loadings from principal components analysis of M	B = nxm matrix of factor loadings on first n variables (which, in the case of the earlier method, means all the variables) for first m principal components derived in a principal components analysis of matrix M.	
4. Diagonal matrix used for scaling, to get 1's in diagonal of final matrix	D = nxn diagonal matrix of terms from diagonal of BB´	
5. Corrected (consistent) matrix	Earlier method $R_{cs} = D^{-\frac{1}{2}} BB' D^{-\frac{1}{2}}$ R_{cs} is the nxn corrected correlation matrix. It is a consistent singular matrix of rank m´.	Revised method $R_c = D^{-\frac{1}{2}} BB' D^{-\frac{1}{2}}$ R_c is the nxn corrected correlation matrix. It is a consistent matrix and its rank* will usually be n, making it a non-singular matrix, though in the special case where m´<m it would probably turn out to be a singular matrix* with rank n − m + m´

*In the extremely unlikely event that R_c turns out ot have a rank lower than n − m + m´ (an eventuality which, however improbable it is, is still mathematically possible since the rank might go as low as m´ − k) additional dummy variables should be added to matrix M, and the entire operation repeated, to obtain a new R_c with acceptable rank (i.e., rank no lower than n − m + m´).

n matrix M

s inconsistent)

eigenvalues > 0
nvalues
ues

f variables, and
ariables added at end of
to produce matrix M.

Revised method

k ≥ n´− m

on first n variables (which,
nod, means all the variables)
ts derived in a principal
M.

om diagonal of BB´

ethod
BB´ D^-½
nxn corrected correlation
. It is a consistent matrix
s rank* will usually be n,
it a non-singular matrix,
in the special case where m´<m
ld probably turn out to be a
ar matrix* with rank n − m + m´
ank lower than n − m + m´ (an
tically possible since the rank
be added to matrix M, and the
rank (i.e., rank no lower

any linear dependency artificially established by the correction-for-inconsistency procedure is likely to derive its character from capitalization on chance and therefore to involve every variable in the matrix. This means that the elimination of any single variable, no matter how trivial, breaks the chain and destroys the dependency. It is at this point that the modification of the original procedure enters the picture.

3. Modification of the earlier procedure

All that is necessary in order to get a corrected matrix of full rank is to include some dummy variables in the original set of raw data on which the correlative matrix is based. These dummy variables are included in the original matrix and are then eliminated after the correction-for-inconsistency procedure has been applied.

This leaves only two questions: how many dummy variables should be used and what kinds of variables are suitable for the purpose? The answer to the first question is simple; at least as many dummy variables are needed as the largest number of zero eigenvalues that are likely to have been artificially produced by the correction-for-inconsistency procedure. This is equivalent to the largest number of negative eigenvalues there are likely to be in the inconsistent matrix. And how many is that? It is easy enough to do a preliminary analysis to determine the eigenvalues of the consistent matrix you are starting with--the matrix that doesn't yet have any dummy variables added--but how can you tell that will happen to the eigenstructure when you add some unspecified number of variables? The answer to that is that if the dummy variables are of a suitable kind they won't cause any increase in the number of negative eigenvalues. That brings us right back to our second question: what kinds of variables are suitable for use as dummy variables? The answer is simple; ideally they should be variables that don't have any significant correlation with any of the other variables in the matrix, so that they will be unlikely to produce another near-zero eigenvalue which would be in danger of being pushed below zero by whatever the situation is that has caused the matrix to be inconsistent in the first place.

The best dummy variables are probably random numbers, though if this is inconvenient there are lots of alternatives that would probably work out well--for instance an arbitrary serial number that may have been assigned to identify each case; or even, perhaps, a substantive variable that is of little importance and is known to have negligible correlations with just about everything else, so that you don't really want it in your final matrix.

The answer to our two questions about kind and number of dummy variables, then, is that random numbers are probably about the best kind, and that at least as many of these dummy variables should be included as the number of negative eigenvalues the inconsistent matrix has before any dummy variables are added to it. But it doesn't do any harm to play safe by including more dummy variables than are actually needed. And incidentally, it doesn't matter at all whether the random number variables have been normalized or are used in their raw rectangular state.

D. Application of the New Procedure

To test the procedure, it was tried out on a matrix that had deliberately been made inconsistent by using underestimates of the reliabilities to correct for attenuation. KR21 reliability coefficients were used for the purpose. This is something that might easily occur in an operational situation--if, for instance, KR21 reliabilities, which are known to be systematically too low in most circumstances, were the only kind available.

Ten test variables from the TALENT battery were included and four random-number variables were added. This was about three more dummy variables than seemed likely to be needed, but extra dummy variables do no harm. The results are presented in Tables 2 and 3. The following is notation used in those two tables:

R is the raw score correlation matrix among 14 variables (10 TALENT tests and 4 random-number variables).

$$n' = 14$$
$$n = 10$$
$$k = 4$$

M is a 14 x 14 inconsistent matrix produced by overcorrecting matrix R for attenuation. KR21's were used for the purpose, except for one test, for which the reliability coefficient (split-half) was based on the experimental form items.

M_{10} is the 10 x 10 inconsistent matrix produced by eliminating the last 4 rows and columns (i.e., the 4 random-number variables) from matrix M.

R_{∞} is the 10 x 10 consistent non-singular correlation matrix produced by using more accurate reliability estimates to correct matrix R for attenuation than were used to produce matrix M. Split-half reliabilities corrected by Angoff formula 16 were used for the purpose, in obtaining R_{∞}.

R_c is the 10 x 10 consistent non-singular correlation matrix obtained by correcting matrix M for inconsistency, using the revised procedure.

The original 14-variable correlation matrix, R, before correction for attenuation, is shown in the upper right half of Table 2. The lower left half of this table contains matrix M, the matrix obtained by using KR21 reliabilities to correct for attenuation. Although the KR21's didn't underestimate test reliability badly enough to make any of the correlations corrected for attenuation go over 1, they were enough too low to make the matrix inconsistent. The initial test for consistency, based just on the 10-variable matrix (M_{10}), showed one negative eigenvalue. The full M matrix, with 14 variables, also had just one negative eigenvalue, which supports the idea that adding random-number variables to an inconsistent matrix isn't likely to increase the number of negative eigenvalues. The lower left half of Table 3 contains the corrected matrix, R_c, which is consistent and nonsingular.

To get some idea of what the final matrix would have looked like if better reliability estimates had been used initially, another matrix was computed, using split-half reliabilities instead of KR21's to correct

Table 2. Creation of an inconsistent correlation matrix M (by overcorrection for attenuation*)

First 10 variables: TALENT test scores
Last 4 variables: Random numbers

Cases: A representative 10% sample of those 12th-grade boys in Project TALENT having scores on all 10 of the test variables.

No. of cases: For the original correlation matrix (R), N=3689 12th grade boys

	TALENT test		KR 21** Reliab.	Original matrix R (upper right) and inconsistent matrix M (lower left)													
				R-102 1	R-103 2	R-106 3	R-107 4	R-112 5	R-230 6	R-250 7	R-282 8	R-290 9	R-312 10	X1 11	X2 12	X3 13	X4 14
1	R-102	Vocab.	.700		.726	.690	.733	.568	.614	.723	.417	.506	.615	-.014	.012	-.022	-.003
2	R-103	Lit Inf	.766	.992		.648	.672	.398	.582	.700	.321	.442	.575	-.020	.012	-.024	.020
3	R-106	Math Inf	.886	.876	.787		.764	.405	.608	.639	.456	.554	.840	-.026	.024	-.021	.004
4	R-107	PhySciInf	.814	.970	.851	.900		.515	.558	.650	.442	.521	.689	-.008	.022	-.026	-.005
5	R-112	Mech Inf	.677	.826	.552	.523	.694		.345	.441	.400	.341	.349	-.006	.010	-.004	-.003
6	R-230	Eng	.896	.775	.703	.682	.654	.443		.659	.366	.501	.653	-.029	-.003	-.021	.005
7	R-250	Rdg	.859***	.932	.863	.732	.777	.578	.751		.440	.577	.616	-.029	.005	-.023	.020
8	R-282	Vis 3-D	.721	.586	.432	.570	.577	.573	.456	.559		.579	.442	-.026	.015	-.009	-.001
9	R-290	Abs Reas	.655	.747	.624	.527	.713	.512	.654	.770	.842		.558	-.027	.018	.004	.002
10	R-312	Math II	.846	.800	.714	.971	.831	.462	.750	.722	.565	.750		-.029	.029	-.011	-.006
11	X_1	Random	1.000	-.016	-.023	-.027	-.009	-.008	-.031	-.031	-.030	-.034	-.031		.001	-.009	-.028
12	X_2	Random	1.000	.014	.013	.025	.025	.012	-.004	.005	.018	.022	.032	.001		.019	-.005
13	X_3	Random	1.000	-.026	-.028	-.023	-.028	-.005	-.022	-.025	-.011	.005	-.012	-.009	.019		.010
14	X_4	Random	1.000	-.003	.023	.004	-.006	-.004	.005	.021	-.001	.002	-.007	-.028	-.005	.010	
	Eigenvalues																
	Matrix M_{10}			7.43	.86	.72	.47	.32	.13	.09	.05	.01	-.08				
	Matrix M			7.43	1.04	1.02	.98	.97	.85	.72	.46	.32	.13	.09	.06	.01	-.08

* Matrix M was produced by using reliability coefficients that are systematic underestimates, in correcting the original matrix (R) for attenuation.

** KR 21 except for Reading Comp. Test (R-250). These reliability coefficients come from Table 4-7 (cols. 9 and 12) of:

Shaycoft, Marion F. The high school years: Growth in cognitive skills.
Project TALENT Office, American Institutes for Research and Univ. of Pittsburgh, 1967

Reliability coefficients for the four random number variables were set at 1.000 arbitrarily, which is equivalent to not correcting for attenuation on those variables.

*** This is a split-half reliability estimate based on the experimental form of the test and corrected for length by the Spearman-Brown formula; the split was on the basis of passages (as in the corresponding split-half reliability coefficient in Table 3) rather than individual items.

Table 3. Consistent matrix R_c (obtained by correcting matrix M for inconsistency) and initially consistent matrix R_∞, both matrices showing correlations corrected for attenuation*

	TALENT test	Split-half reliab.*	Matrices R_c (lower left) and R_∞ (upper right)									
			R-102 1	R-103 2	R-106 3	R-107 4	R-112 5	R-230 6	R-250 7	R-282 8	R-290 9	R-312 10
1.	R-102 Vocab	.778		.911	.828	.902	.752	.725	.855	.529	.665	.738
2.	R-103 Lit Inf	.817	.935		.760	.807	.513	.671	.808	.399	.566	.673
3.	R-106 Math Inf	.892	.841	.786		.878	.500	.671	.705	.541	.680	.942
4.	R-107 PhySci Inf	.848	.932	.850	.900		.652	.632	.736	.538	.656	.792
5.	R-112 Mech Inf	.735	.781	.555	.524	.694		.419	.536	.523	.461	.431
6.	R-230 Eng	.921	.747	.701	.683	.654	.443		.716	.428	.605	.720
7.	R-250 Rdg	.919	.895	.862	.732	.778	.579	.751		.514	.698	.680
8.	R-282 Vis 3-D	.796	.572	.429	.569	.576	.570	.455	.558		.752	.524
9.	R-290 Abs Reas	.744	.723	.622	.727	.713	.512	.654	.770	.842		.685
10.	R-312 Math II	.893	.780	.709	.970	.829	.459	.750	.721	.566	.749	
	Eigenvalues											
	For matrix R_c		7.36	.85	.72	.47	.32	.13	.09	.05	.01	.005-
	For matrix R_∞		7.02	.87	.74	.49	.36	.20	.16	.11	.04	.01

* The reliability coefficients come from Table 4-7, col. 5 of (Shaycoft, op. cit.)

the original matrix, R, for attenuation. This new matrix, R_∞, turned out to be consistent and non-singular. It is shown in the upper right half of Table 3. Matrix R_c is much closer to this "correct" matrix, R_∞, than the inconsistent matrix, M_{10}, is.

In summary, then, in converting an inconsistent correlation matrix to a consistent one it is possible through the use of dummy variables to maintain the full rank of the original matrix, instead of converting a non-singular matrix into a singular one, as occurs when dummy variables aren't used.