IMAGE ANALYSIS FOR SINGULAR DATA

Mike Pore, University of New Orleans

Image analysis was presented by Guttman [1953], and two adaptations to factor analysis have been made. The first is an obvious, natural technique involving the same principle as principle factor analysis. The second was formulated by Harris [1963] along the same basic lines used in cononical factor analysis.

Definition. Image analysis is the partitioning of the observable random vector, X into an image vector, WX, and an anti-image vector, (I - W)X.

$$X = WX + (I - W)X$$

where WX is the minimum squared error linear estimate of X such that

$$\text{diag } W = \phi, \ \phi = \text{the null matrix.}$$

This is a "natural" partitioning of X that separates that part of X not explained by $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_N$ from that part which is explained by the other N - 1 variables, the anti-image, corresponds with the factor analysis concept of unique variables, but does not correspond to the way the model characterized the unique variables as being uncorrelated to each other and to each factor, hence

$$E(U) = \phi$$

$$\text{Cov}(Y,U) = \phi$$

$$\text{Cov}(U,U) = D, \text{ a diagonal matrix.}$$

Guttman's discovery (that this natural partitioning into image and anti-image existed) sheds new light on how the unique variables of factor analysis should be modeled. This will be discussed later. Following are some of the results of the above definition.

Theorem. The coefficient matrix W in the image of X is a solution to the equations

$$\text{diag } W = \phi$$

$$WR = R - D$$

for some diagonal matrix D.

Proof. Presented by Guttman [1953]. His development does not require that R be non-singular, however, the development beyond this theorem usually does. Hence, this is where this development diverges from the classical development. Some properties of image analysis can be derived from this theorem, using the Moore-Penrose "pseudoinverse" of a matrix.

Definition and Theorem. For each matrix A there exists a unique matrix A^+ (called the pseudoinverse of A) such that

(1) $AA^+A = A$

(2) $A^+AA^+ = A^+$

(3) $(AA^+)^T = AA^+$

(4) $(A^+A)^T = A^+A$

Property. If X is an unknown matrix satisfying the linear system

$$XA = B$$

for constant matrices A and B, then

(1) the system is consistent if and only if $BA^+A = B$ and

(2) $X = BA^+A + Z(I - AA^+)$ for some matrix Z. Furthermore,

(3) $X = BA^+A + Z(I - AA^+)$ satisfies $XA = B$ for each matrix Z of appropriate dimensions.

A complete development of pseudoinverses can be found in Lewis and Odell [1971], Boullion and Odell [1971], or Graybill [1969].

Property. The covariance of the image vector is

$$\text{Cov}(WX,WX) = W\ R\ W^T = R + DR^+D - 2D.$$

Proof. Assume WR = R - D is consistent, i.e. $DR^+R = D$. R can be factored as

$$R = Q_1^T\ [\lambda_i]\ Q_1$$

where $[\lambda_i]$ is a full rank diagonal matrix with the positive eigenvalues of R on the main diagonal. Q_1 is a matrix of eigenvectors respective to the eigenvalues of $[\lambda_i]$. The result may easily be obtained by matrix manipulation and use of the two equations:

$$Q_1\ Q_1^T = I$$

$$R^+ = Q_1^T\ [\lambda_i]^{-1} Q_1 \quad .$$

Definition from factor analysis. The diagonal elements of $W\ R\ W^T$ are called the communalities of X. This definition of communalities is also in Horst [1965] and Rummell [1970].

Property. The communalities of image analysis are

$$\text{communalities} = \text{diag } [W\ R\ W^T] = I - D$$

Proof.

$$R - D = W\ R_T$$
$$= R\ W^T$$

so

$$W\ R\ W^T = W\ (R - D)$$
$$= W\ R - W\ D$$
$$= R - D - W\ D.$$

Hence

$$W\ D = R - D - W\ R\ W^T$$
$$= D - D\ R^+D$$

by the above Property. But diag W = ϕ, so

$$\phi = \text{diag } W D$$
$$= D - \text{diag } D R^{+} D.$$

This, then, yields the result.

There is no general solution to Guttman's theorem in the literature. The solution for R non-singular is widely known, and can be found in Harris [1962, 1963], Guttman [1954], or Rummell [1970]. But no one has presented a solution for singular R, and in fact, there are erroneous solutions in circulation (see Fisher [1970]). The following development will satisfy this need for a solution, and indications are made for deriving other, equivalent solutions.

Lemma. There are no zeros on the main diagonal of R^{+}.

Proof. Let r = rank R, then

$$R = Q[\lambda_i]\, Q^T$$

where $[\lambda_i]$ is an $r \times r$ diagonal matrix of positive eigenvalues of R, and Q is an $N \times r$ matrix of respective eigenvectors. Now

$$R^{+} = Q[\lambda_i]^{-1} Q^T.$$

This may be verified by substituting R and R^{+} into the defining equations of pseudoinverse. Now observe that the $i^{\underline{th}}$ diagonal element of R is

$$1 = \sum_{j=1}^{r} (q_{ij}^2\ \lambda_i)$$

and each term in the sum is non-negative. Therefore each term in the sum

$$\sum_{j=1}^{r} \frac{q_{ij}^2}{\lambda_i}$$

is also non-negative and the sum is not zero. Examine the $i^{\underline{th}}$ diagonal element of R^{+}. It is given by the sum above.

Lemma. $D(I - R^{+}R) = \phi$, if and only if

$$D \text{ diag } (I - R^{+}R) = \phi .$$

Proof. Follows from $I - R^{+}R$ being symmetric and idempotent.

Notation. Let $A = \text{diag } (I - R^{+}R) = (I - \text{diag } R^{+}R)$.

Theorem. $D R^{+}R = D$ if and only if D is of the form $D = H(I - AA^{+})$ for some $N \times N$ arbitrary matrix H.

Proof. Rather than find necessary and sufficient conditions that $D R^{+}R = D$, necessary and sufficient and conditions are found that

$$D A = \phi.$$

The lemma establishes the equivalence of the two. The result then follows immediately.

A is a diagonal matrix, hence to impose the restriction that D be diagonal, we need only require that H in the theorem above be diagonal. This represents only admissible D's, not solutions to the original problem.

A solution is now given to the first theorem.

Theorem. A solution to the system of equations

$$W R = R - D$$
$$D \text{ is diagonal}$$
$$\text{diag } W = \phi$$

is

$$W_0 = I - (I - AA^{+})(\text{diag } R^{+})^{-1}R^{+} - A^{+}(I - R R^{+})$$

and

$$D_0 = (I - AA^{+})(\text{diag } R^{+})^{-1}.$$

Proof. Since A is diagonal, D_0 is diagonal. Taking the diagonal of W_0 it easily collapses to the null matrix, ϕ. Postmultiplication of W_0 by R yields

$$W_0 R = R - (\text{diag } R^{+})^{-1}(I - AA^{+})R^{+}R.$$

since $D_0\ A = \phi$ then $D_0\ (I - R^{+}R) = \phi$ by the second lemma, hence

$$W_0 R = R - (\text{diag } R^{+})^{-1}(I - A\ A^{+}) = R - D$$

since diagonal matrices commute under multiplication. In general, all solutions to Guttman's theorem are of the form

$$W = I - H_1\ (I - A\ A^{+})R^{+} + H_2(I - R\ R^{+})$$

for some H_1 and H_2 where H_1 is diagonal. This theorem presents the solution when H_2 is also diagonal.

Corollary. In the theorem above, if R is non-singular, then the unique solution is

$$W_0 = I - (\text{diag } R^{-1})^{-1}R^{-1}$$

and

$$D_0 = (\text{diag } R^{-1})^{-1}.$$

Proof. A proof of this is presented with each description of image analysis in the literature, but can also be seen to follow from this theorem by observing that $A = \phi$.

The following two examples demonstrate the "factoring" process involved in image analysis, and some of the problems encountered when using it in any data compression technique. Consider the two (contrived) correlation matrices, R_1 and R_2:

$$R_1 = \begin{matrix} 1.00 & .50 & .50 & .25 & .25 \\ .50 & 1.00 & .50 & .25 & .25 \\ .50 & .50 & 1.00 & .25 & .25 \\ .25 & .25 & .25 & 1.00 & .50 \\ .25 & .25 & .25 & .50 & 1.00 \end{matrix}$$

$$R_2 = \begin{matrix} 1.00 & .50 & .50 & .25 & .25 & .25 \\ .50 & 1.00 & .50 & .25 & .25 & .25 \\ .50 & .50 & 1.00 & .25 & .25 & .25 \\ .25 & .25 & .25 & 1.00 & .50 & .50 \\ .25 & .25 & .25 & .50 & 1.00 & 1.00 \\ .25 & .25 & .25 & .50 & 1.00 & 1.00 \end{matrix}$$

R_1 is a full rank (rank = 5) matrix presumably representing five variables, say $(x_1, x_2, x_3, x_4, x_5)$. R_2 is a singular (rank = 5) matrix representing the same five variables as R_1, but with a sixth variable identical to the fifth: $x_6 = x_5$. The resulting image covariance matrices, G_1 and G_2, were calculated using the last theorem:

$$G_1 = \begin{matrix} .244 & .295 & .295 & .205 & .205 \\ .295 & .344 & .295 & .205 & .205 \\ .295 & .295 & .244 & .205 & .205 \\ .205 & .205 & .205 & .276 & .175 \\ .205 & .205 & .205 & .175 & .175 \\ .205 & .205 & .205 & .175 & .276 \end{matrix}$$

$$G_2 = \begin{matrix} .344 & .295 & .295 & .205 & .250 & .250 \\ .295 & .244 & .295 & .205 & .250 & .250 \\ .295 & .295 & .344 & .205 & .250 & .250 \\ .205 & .205 & .205 & .276 & .500 & .500 \\ .250 & .250 & .250 & .500 & 1.000 & 1.000 \\ .250 & .250 & .250 & .500 & 1.000 & 1.000 \end{matrix}$$

Notice that the fifth row and column of G_1 and G_2 differ (although they partition the same variable: x_5) and the last two rows and columns of G_2 are identical to R_2.

The anti-image covariance matrices also show this by the zeros in $R_2 - G_2$:

$$R_1 - G_1 = \begin{matrix} .656 & .205 & .205 & .045 & .045 \\ .205 & .656 & .205 & .045 & .045 \\ .205 & .205 & .656 & .045 & .045 \\ .045 & .045 & .045 & .724 & .325 \\ .045 & .045 & .045 & .325 & .724 \end{matrix}$$

$$R_2 - G_2 = \begin{matrix} .656 & .205 & .205 & .045 & 0 & 0 \\ .205 & .656 & .205 & .045 & 0 & 0 \\ .205 & .205 & .656 & .045 & 0 & 0 \\ .045 & .045 & .045 & .724 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

These examples demonstrate that, while the "filtering" process of image analysis $(X = WX + (I-W)X)$ does partition the unique part of each variable from the linearly predictable part (in terms of the other variables), this partitioning is not necessarily a partitioning of the relevant from the irrelevant. Those variables that are linear combinations of the others are reproduced in the image in their entirety, while data compression techniques seek to eliminate such variables in order to reduce redundancy.

Some data compression techniques, such as image covariance factor analysis, use image analysis as a conditioning filter to partition relevant and irrelevant parts of each variable. The failure of image analysis to exclude redundant variables may cause the factor analysis to deal with some of the variables differently for the non-singular and singular cases: e.g. consider x_5 in the examples above.

In summary, image analysis was developed for singular correlation matrices, and an example of how variable dependence effects image analysis was presented. The author has also developed a mathematically consistent factor analysis model employing image analysis [Pore, 1973, 1974] in addition to the two mentioned in the first paragraph of this paper. The examples indicate, however, that if a data compression technique is intended (as in factor analysis), then great care is recommended in interpreting the results of pre-conditioning or filtering singular data with image analysis.

REFERENCES

Boullion, Thomas L., and Patrick L. Odell (1971). Generalized Inverse Matrices. New York: John Wiley & Sons, Inc.

Fisher, David R. (1970). A Comparison of the Various Techniques of Multiple Factor Analysis Applied to Biosystematic Data. Ph.D. Dissertation, University of Kansas.

Graybill, Franklin A. (1969). Introduction to Matrices With Applications In Statistics. Belmont, California: Wadsworth Publishing Co., Inc.

Guttman, L. (1953). Image Theory for the Structure of Quantitative Variates. Psychometrika, 18, 277-96.

________ (1954). Some Necessary Conditions for Common Factor Analysis. Psychometrika, 19, 149-61.

Harris, C.W. (1962). Some Rao-Guttman Relationships. Psychometrika, 27, 247-63.

________ (1963). Canonical Factor Models for the Description of Change. Problems in Measuring Change.

Madison, Wisconsin: University of Wisconsin Press.

Horst, Paul (1965). Factor Analysis of Data Matrices. New York: Holt, Rinehart and Winston, Inc.

Lewis. T.O., and P.L. Odell (1971). Estimation in Linear Models. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Pore, Michael D. (1973). Factor Analysis as a Data Compression Technique. Ph.D. Dissertation, Texas Tech University.

________(1974). Linear Compressions Using the Image Factor Analysis Criterion. Presented at the Annual Meeting of the American Statistical Association. Published in the Proceedings of the Social Statistics Section. 400-3.

Rummel, Rudolph J. (1970). Applied Factor Analysis. Evanston, Illinois: Northwestern University Press.