# Ill-conditioned correlation analyzer

## WILLIAM M. HOLMES
*Northeastern University, Boston, Massachusetts 02115*

The ill-conditioned correlation analyzer (ILK) is a set of procedures for detection of ill-conditioning in correlation matrices. The philosophy of the program is to provide a variety of statistics on the internal consistency, conditioning, interdependence, and independence of items in a correlation matrix. The procedures assess presence, strength, and source of ill-conditioning. The program can be run interactively or as a batch process.

**Input.** The program analyzes a correlation matrix. The matrix may be in a file generated by some other program or may be supplied at the time of execution of the program. The sample size on which the correlations are based is also required. Options allow selecting a subset of variables for analysis to more clearly specify sources of ill-conditioning.

**Output.** The program prints out the correlation matrix input, the subset of correlations to be analyzed, and summary measures of ill-conditioning, including the determinant of the matrix, Kuder-Richardson's alpha (Bohrnstedt, 1970; Holmes, 1979), the condition index (Belsley, Kuh, & Welsh, 1980), Bartlett's (1950) test of item independence, Haitovsky's test of multicollinearity (Rockwell, 1975), and F tests of dependence and independence. It also provides statistics for each variable, including the multiple R squared of each variable with all others (Farrar & Glauber, 1967), an F test of the multiple R squared, item tolerances, average correlation of each variable with all others, and the condition number of each item. In addition, the program allows an option of printing the inverse of the matrix analyzed and rerunning the analysis while adding or deleting variables from the subset selected. Example output is given in Table 1.

**Software.** The program is written in FORTRAN. A maximum of 16 variables can be analyzed under the default space allocation. It calls external IMSL subroutines to invert the matrix and evaluate significance probabilities of the F tests and chi squares, although other subroutines can be substituted for them. The program also calls an external function to compute the determinant of the correlation matrix, a copy of which can be supplied if needed. Computation of the matrix inverse and determinant are in double precision.

**Hardware.** The program was developed on a VAX-11/780 operating under VMS. Approximately 50 KB are required for execution. Space requirements can be adjusted by redimensioning the matrices and arrays.

## REFERENCES

BARTLETT, M. S. Tests of significance in factor analysis. *British Journal of Psychology, Statistics Section,* 1950, **3**, 77-85.

BELSLEY, D. A., KUH, E., & WELSH, R. E. *Regression diagnostics.* New York: Wiley, 1980.

BOHRNSTEDT, G. A quick method for determining the reliability and validity of multiple item scales. *American Sociological Review,* 1970, **34**, 542-548.

FARRAR, D., & GLAUBER, R. Multicollinearity in regression analysis. *Review of Economics and Statistics,* 1967, **49**, 92-107.

HOLMES, W. Examining reliability and multicollinearity of scale items. *Behavior Research Methods & Instrumentation,* 1979, **11**, 86.

ROCKWELL, R. C. Assessment of multicollinearity: The Haitovsky test of the determinant. *Sociological Methods and Research,* 1975, **3**, 308-320.

**Table 1**
**Example Output**

SELECTED CORS

```
 1.00 -0.62 -0.58
-0.62  1.00  0.72
-0.58  0.72  1.00
```

SUMMARY RESULTS

ALPHA=−0.697266
DETERMINANT= 0.276221    CONDITION INDEX= 1.38

| F | DF | PROB | BARTLETT | CHIDF | PROB | HAITOVSKY | PROB |
|---|---|---|---|---|---|---|---|
| 104.81 | 40 | 0.00000 | 54.04 | 3. | 0.00000 | 13.58 | 0.00354 |

VARIABLE ANALYSIS

| VARIABLE | R-SQUARED | TOLERANCE | AVG COR | CONDTN | F | PROB |
|---|---|---|---|---|---|---|
| 1 | 0.41949 | 0.5805 | −0.6000 | 1.3848 | 28.90 | 0.0000 |
| 2 | 0.58081 | 0.4192 | 0.0540 | 1.0000 | 55.43 | 0.0000 |
| 3 | 0.55487 | 0.4451 | 0.0700 | 1.0619 | 49.86 | 0.0000 |