Title
Positive Definiteness via Offdiagonal Scaling of a Symmetric Indefinite Matrix

Permalink
https://escholarship.org/uc/item/3cz1w4zj

Authors
Bentler, P M
Yuan, Ke-Hai

Publication Date
2010-02-11

February 11, 2010

Positive Definiteness via Offdiagonal Scaling of a Symmetric Indefinite Matrix

Peter M. Bentler

University of California, Los Angeles

Ke-Hai Yuan

University of Notre Dame

Positive Definiteness via Offdiagonal Scaling of a Symmetric Indefinite Matrix

Abstract

Indefinite symmetric matrices that are estimates of positive definite population matrices occur in a variety of contexts such as correlation matrices computed from pairwise present missing data and multinormal based theory for discretized variables. This note describes a methodology for scaling selected off-diagonal rows and columns of such a matrix to achieve positive definiteness. As a contrast to recently developed ridge procedures, the proposed method does not need variables to contain measurement errors. When minimum trace factor analysis is used to implement the theory, only correlations that are associated with Heywood cases are shrunk.

Positive Definiteness via Offdiagonal Scaling of a Symmetric Indefinite Matrix

Let R be a symmetric indefinite matrix, that is, a matrix with both positive and negative eigenvalues.

Often such matrices are intended to estimate a positive definite (pd) matrix, as can be seen in a wide

variety of psychometric applications including correlation matrices estimated from pairwise or binary

information (e.g., Wothke, 1993). Approaches to modifying R to create a pd matrix for further analysis

include least squares approximation (Knol & ten Berge, 1989) and adding a small constant to its diagonal

(e.g., Yuan & Chan, 2008); a thorough review is given in Yuan, Wu, and Bentler (2009). This note

describes a methodology for scaling off-diagonal elements of R to achieve a pd approximation. This is

done by finding a bounded diagonal scaling matrix that shrinks selected off-diagonal rows and columns

of R.

Lemma 1. There exists a diagonal matrix D with nonzero elements such that $(R - D)$ is positive

semidefinite.

Proof: Such a D can be obtained e.g., by minimum trace factor analysis (Bentler, 1972; Della Riccia &

Shapiro, 1982).

In the standard factor analytic situation where R is positive definite, D would be a pd diagonal matrix of

unique variances, and $(R - D) = FF'$ would be the covariance matrix of the common parts of the

variables. However, in the current context, R is indefinite and hence D has different properties.

Lemma 2. One or more diagonal entries in D are negative.

Proof: Assume the contrary. Then R is the sum of a positive semidefinite (psd) and a pd diagonal matrix,

and thus R would be pd, which is contrary to assumption. Hence D must have one or more negative

diagonal elements.

Let H^2 be a diagonal matrix containing the diagonal of $(R - D)$; in standard factor analysis, the elements of this matrix are known as communalities. Let D_R be the diagonal matrix containing the diagonal of R, and let $R_0 = (R - D_R)$. With these definitions, $R - D = R_0 + H^2$. Let Δ be a pd diagonal matrix such that $0 < \Delta^2 H^2 < D_R$.

Theorem. $R^* = \Delta R_0 \Delta + D_R$ is positive definite.

Proof: Note that $\Delta R_0 \Delta + \Delta^2 H^2 = \Delta(R_0 + H^2)\Delta$ is psd and $D_R - \Delta^2 H^2$ is pd. Since

$R^* = (\Delta R_0 \Delta + \Delta^2 H^2) + (D_R - \Delta^2 H^2)$ is the sum of a psd and a pd matrix, it is pd.

The theorem shows how to obtain a pd matrix from an indefinite one, where $diag(R^*) = diag(R)$ and the offdiagonals of R^* are rescaled elements of R. When R is a correlation matrix with unit diagonals, R^* can be similarly interpreted.

Application

Suppose that R is a correlation matrix obtained by polychoric and/or polyserial methodology. It is well known that this matrix is often indefinite in small samples, leading to problems in estimation and testing of derived models such as structural equation models. R^* may be an appropriate substitute for R in such models. Although biased, R^* is a consistent estimate of the population counterpart to R since R^* approaches R as N goes to infinity. The sampling distribution of elements of R^* can be obtained using the bootstrap.

To obtain R^* in practice, a minimum trace factor analysis algorithm (e.g., Bentler, 1972; Jamshidian & Bentler, 1998) applied to R will yield a unique H^2 such that $tr(H^2)$ is minimized. Let H_i^2 be the ith diagonal element of H^2. If R is indefinite, many elements will have $H_i^2 < 1$ but one or more elements

will be Heywood cases with $H_i^2 \geq 1$. The matrix Δ^2 is constructed such that an element $\Delta_i^2 = 1$ if

$H_i^2 < 1$, while if $H_i^2 \geq 1$, $\Delta_i^2 = k / H_i^2$ for some a priori constant $k < 1$. For simplicity, k may be taken as

the same value for all Heywood variables. It is desirable to have k be only marginally smaller than 1.0,

for example, $k = .96$. Then if $H_i^2 = 1.1$, for example, the ith row and column of R_0 is multiplied by .934

to yield the correlation in $R*$. Only those variables corresponding to Heywood cases have their

correlations rescaled; the remainder are not modified.

An example of this methodology is given in Table 1, which shows the correlations among 12 variables

obtained for a random sample of 50 cases from a categorized multinormal population, based on Bonett

and Price's (2005) odds-ratio tetrachoric estimator. The eigenvalues of this correlation matrix are

6.4233, 1.3704, 1.1237, 0.7641, 0.7174, 0.5059, 0.4430, 0.3334, 0.1559, 0.1115, 0.0600, -0.0087. The

small negative eigenvalue makes the matrix indefinite. Minimum trace factor analysis showed that

variables 3, 4, 6, and 9 had communalities greater than 1.0, ranging from 1.037 to 1.1543. Table 2 gives

the correlation matrix after scaling using $k = .96$. Only variables 3, 4, 6, and 9 have correlations that are

reduced. The median reduction in correlation is .027, while the maximum reduction is .0712 (r_{43} reduced

from .8579 to .7867). The eigenvalues of the resulting matrix are 6.2305, 1.3369, 1.1195, 0.7738, 0.7204,

0.5146, 0.4473, 0.3641, 0.1780, 0.1226, 0.1181, 0.0742.

Discussion

The most widely known methodology for dealing with indefinite or near singular symmetric matrices is

that of ridge regression (Hoerl & Kennard, 1970) or Tikhonov regularization[1]. In standard ridge

regression and other ridge applications, each diagonal of a symmetric matrix is incremented by a small

positive number, say κ, that is larger than the smallest eigenvalue of R. The statistical theory to make

---

[1] http://en.wikipedia.org/wiki/Tikhonov_regularization

this approach well-rationalized in the context of covariance and correlation structures has recently been developed (e.g., Yuan & Chan, 2008; Yuan, Wu, & Bentler, 2009), where variables are assumed to contain measurement errors that are explicitly accounted for in the model. The approach proposed in this paper does not need variables to contain measurement errors in application. An example is the regression model with standardized variables when the correlation matrix of the predictors is nonpositive definite. When the proposed procedure is implemented using minimum trace factor analysis, correlations for variables associated with Heywood cases are smoothly scaled down; those among non-Heywood variables remain undisturbed.

A limitation of this methodology is that the scaling constant k is subjectively determined. The example used $k = .96$, but other values marginally below 1.0 could be used as well. Limited experience shows that the precise value does not matter much. The need to use subjective judgment in selecting tuning values is also a feature of previously proposed methods (Knol & ten Berge, 1989; Yuan & Chan, 2008; Yuan, Wu, & Bentler, 2009).

References

Bentler, P. M. (1972). A lower bound method for the dimension-free measurement of internal

consistency. *Social Science Research*, *1*, 343-357.

Bonett, D. G., & Price, R. M. (2005). Inferential methods for the tetrachoric correlation coefficient.

Journal of Educational and Behavioral Statistics, 30, 213-225.

Della Riccia, G., & Shapiro, A. (1982). Minimum rank and minimum trace of covariance matrices.

Psychometrika, *47*, 443-448.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems.

Technometrics, *12*, 55–67.

Jamshidian, M., & Bentler, P. M. (1998). A quasi-Newton method for minimum trace factor analysis.

Journal of Statistical Computation and Simulation, *62*, 73-89.

Knol, D. L., & ten Berge, J. M. F. (1989). Least squares approximation of an improper correlation matrix

by a proper one. *Psychometrika*, *54*, 53–61.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen & J. S. Long

(Eds.), *Testing structural equation models* (pp. 256–293). Newbury Park, CA: Sage.

Yuan, K.-H., & Chan, W. (2008). Structural equation modeling with near singular covariance matrices.

Computational Statistics & Data Analysis, *52*, 4842–4858.

Yuan, K.-H., Wu, R., & Bentler, P. M. (2009). Ridge structural equation modeling with correlation

matrices for ordinal and continuous data. *British Journal of Mathematical and Statistical Psychology*.

Table 1

Tetrachoric Correlations (Bonett-Price Estimator)

1.0000	0.2387	0.6161	0.6167	0.6621	0.5173	0.6758	0.7071	0.7983	0.5769	0.4705	0.7881
0.2387	1.0000	0.3506	0.3537	0.2959	0.4637	0.1931	0.1202	0.2316	0.1708	0.4047	0.1161
0.6161	0.3506	1.0000	0.8579	0.6603	0.4093	0.3826	0.5164	0.6079	0.5574	0.4512	0.5128
0.6167	0.3537	0.8579	1.0000	0.7477	0.1803	0.4705	0.6167	0.6218	0.4705	0.3582	0.2966
0.6621	0.2959	0.6603	0.7477	1.0000	0.3537	0.7364	0.5670	0.6613	0.5140	0.5140	0.4610
0.5173	0.4637	0.4093	0.1803	0.3537	1.0000	0.3582	0.1803	0.0605	0.4705	0.3582	0.6161
0.6758	0.1931	0.3826	0.4705	0.7364	0.3582	1.0000	0.4705	0.6424	0.6090	0.4911	0.4962
0.7071	0.1202	0.5164	0.6167	0.5670	0.1803	0.4705	1.0000	0.7149	0.4705	0.3582	0.5164
0.7983	0.2316	0.6079	0.6218	0.6613	0.0605	0.6424	0.7149	1.0000	0.4371	0.4371	0.6079
0.5769	0.1708	0.5574	0.4705	0.5140	0.4705	0.6090	0.4705	0.4371	1.0000	0.3745	0.4512
0.4705	0.4047	0.4512	0.3582	0.5140	0.3582	0.4911	0.3582	0.4371	0.3745	1.0000	0.4512
0.7881	0.1161	0.5128	0.2966	0.4610	0.6161	0.4962	0.5164	0.6079	0.4512	0.4512	1.0000

Table 2

Scaled Correlations

1.0000	0.2387	0.5928	0.5878	0.6621	0.4719	0.6758	0.7071	0.7511	0.5769	0.4705	0.7881
0.2387	1.0000	0.3373	0.3371	0.2959	0.4229	0.1931	0.1202	0.2179	0.1708	0.4047	0.1161
0.5928	0.3373	1.0000	0.7867	0.6353	0.3593	0.3682	0.4968	0.5503	0.5363	0.4341	0.4934
0.5878	0.3371	0.7867	1.0000	0.7127	0.1567	0.4485	0.5878	0.5577	0.4485	0.3414	0.2827
0.6621	0.2959	0.6353	0.7127	1.0000	0.3226	0.7364	0.5670	0.6222	0.5140	0.5140	0.4610
0.4719	0.4229	0.3593	0.1567	0.3226	1.0000	0.3267	0.1644	0.0519	0.4292	0.3267	0.5620
0.6758	0.1931	0.3682	0.4485	0.7364	0.3267	1.0000	0.4705	0.6044	0.6090	0.4911	0.4962
0.7071	0.1202	0.4968	0.5878	0.5670	0.1644	0.4705	1.0000	0.6726	0.4705	0.3582	0.5164
0.7511	0.2179	0.5503	0.5577	0.6222	0.0519	0.6044	0.6726	1.0000	0.4113	0.4113	0.5719
0.5769	0.1708	0.5363	0.4485	0.5140	0.4292	0.6090	0.4705	0.4113	1.0000	0.3745	0.4512
0.4705	0.4047	0.4341	0.3414	0.5140	0.3267	0.4911	0.3582	0.4113	0.3745	1.0000	0.4512
0.7881	0.1161	0.4934	0.2827	0.4610	0.5620	0.4962	0.5164	0.5719	0.4512	0.4512	1.0000