Higher Rank Approximation and Matrix Decomposition Factor Analysis

Kohei Adachi, Osaka University

1. Introduction

In this paper, theorems are presented for the problem of approximating a given matrix by an orthonormal one with its rank being higher than the former matrix. Then, the theorems are used for forming an algorithm of the matrix decomposition factor analysis (MDFA) that has recently proposed for exploratory factor analysis.

2. Higher Rank Approximation (HRA)

Theorem 1. The solution for the HRA problem

$$\min_{\mathbf{H}} \phi(\mathbf{H}) = ||\mathbf{H} - \mathbf{M}||^2 \text{ subject to } q = \text{rank}(\mathbf{H}) \geq \text{rank}(\mathbf{M}) = r \text{ and } \mathbf{H}'\mathbf{H} = \mathbf{I}_q \quad (1)$$

is solved for $\mathbf{H} = \mathbf{KL}' = \mathbf{K}_1\mathbf{L}_1' + \mathbf{K}_2\mathbf{L}_2'$ with $\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2]$, $\mathbf{L} = [\mathbf{L}_1, \mathbf{L}_2]$, $\mathbf{K}'\mathbf{K} = \mathbf{L}'\mathbf{L} = \mathbf{I}_q$, and $\mathbf{K}_1\mathbf{L}_1'$ is given through the singular value decomposition (SVD) $\mathbf{M} = \mathbf{K}_1\Delta\mathbf{L}_1'$ with Δ an $r \times r$ positive definite diagonal matrix.

Proof. (1) is equivalent to maximizing $\eta(\mathbf{H}) = \text{tr}\mathbf{M}'\mathbf{H} = \text{tr}\mathbf{L}_1\Delta\mathbf{K}_1'\mathbf{H} = \text{tr}\mathbf{K}_1'\mathbf{H}\mathbf{L}_1\Delta$ subject to the same constraints. Since $\mathbf{K}_1'\mathbf{H}\mathbf{L}_1$ is sub-orthonormal, $\eta(\mathbf{H}) \leq \text{tr}\Delta$ (ten Berge, 1993). This upper limit is attained for $\mathbf{H} = \mathbf{KL}'$. □

It is found that $\mathbf{K}_2\mathbf{L}_2'$ and $\mathbf{H}$ are not uniquely determined. However, $\mathbf{H}'\mathbf{M}$ for the HRA problem can be uniquely determined as shown next:

Theorem 2. (1) is attained for

$$\mathbf{M}'\mathbf{H} = \mathbf{L}_1\Delta\mathbf{L}_1' \quad (2)$$

Proof. We can rewrite $\phi(\mathbf{H})$ as $c - 2\text{tr}\mathbf{M}'\mathbf{H}$ with c the constant irrelevant to $\mathbf{H}$, which implies that (1) is attained for obtaining the optimal $\mathbf{M}'\mathbf{H}$. It is found to be given by (2) using $\mathbf{M} = \mathbf{K}_1\Delta\mathbf{L}_1'$ and $\mathbf{H} = \mathbf{KL}'$. □

3. Matrix Decomposition Factor Analysis (MDFA)

For a column-centered n-observations × p-variables data matrix $\mathbf{X}$ with $n > p$, MDFA is formulated as

$$\min_{\mathbf{Z},\mathbf{B}} ||\mathbf{X} - \mathbf{F}\Lambda' - \mathbf{U}\Psi||^2 = ||\mathbf{X} - \mathbf{ZB}'||^2 \text{ subject to } n^{-1}\mathbf{Z}'\mathbf{Z} = \mathbf{I}_{p+m} \text{ and } \Psi \text{ being diagonal} \quad (3)$$

(Adachi, 2012; Unkel & Trendafilov, 2010). Here, $\mathbf{Z} = [\mathbf{F}, \mathbf{U}]$ and $\mathbf{B} = [\Lambda, \Psi]$ are block matrices with $\mathbf{F}$ ($n \times m$), $\mathbf{U}$ ($n \times p$), Λ ($p \times m$), and Ψ ($p \times p$) being common factor, unique factor, loading, and square-root unique variance matrices, respectively. The loss function in (3) can be rewritten as $||\mathbf{Z} - \mathbf{XB}'||^2 + c_1$ and $||\mathbf{B} - n^{-1}\mathbf{X}'\mathbf{Z}||^2 + c_2$ with c_1 the constant irrelevant to $\mathbf{Z}$ and c_2 irrelevant to $\mathbf{B}$. Thus, (3) can be attained by alternately performing

$$\min_{\mathbf{Z}} ||\mathbf{Z} - \mathbf{XB}'||^2 \text{ subject to } n^{-1}\mathbf{Z}'\mathbf{Z} = \mathbf{I}_{p+m} \text{ with } \mathbf{B} \text{ kept fixed,} \quad (4)$$

$$\min_{\mathbf{B}} ||\mathbf{B} - n^{-1}\mathbf{X}'\mathbf{Z}||^2 \text{ subject to } \Psi \text{ being diagonal with } \mathbf{C} \text{ kept fixed.} \quad (5)$$

4. Covariances of Variables to Common and Unique Factors

Problem (4) is an HRA problem, since rank($\mathbf{Z}$) is $p + m$ and rank($\mathbf{XB}$) ≤ min(n, p). Thus, we can set $\mathbf{M}$ and $\mathbf{H}$ in Theorem 1 as $\mathbf{M} = n^{-1/2}\mathbf{XB}$ and $\mathbf{H} = n^{-1/2}\mathbf{Z}$ to find that the solution for (4) is given by

$$\mathbf{Z} = n^{1/2}\mathbf{K}_1\mathbf{L}_1' + n^{1/2}\mathbf{K}_2\mathbf{L}_2' = \mathbf{XBL}_1\Delta^{-1}\mathbf{L}_1' + n^{1/2}\mathbf{K}_2\mathbf{L}_2' \quad (6)$$

with $n^{-1/2}\mathbf{XB} = \mathbf{K}_1\Delta\mathbf{L}_1'$. Although (6) is not unique, Theorem 2 shows that $\mathbf{M}'\mathbf{H} = n^{-1}\mathbf{B}'\mathbf{X}'\mathbf{Z}$ is determined as

$$n^{-1}\mathbf{B}'\mathbf{X}'\mathbf{Z} = \mathbf{L}_1\Delta\mathbf{L}_1'. \quad (7)$$

It is shown next that the covariances of variables to factors can be uniquely determined for rank($\mathbf{XB}$) = p.

Theorem 3. If rank($\mathbf{XB}$) = p, then $n^{-1}\mathbf{X}'\mathbf{Z}$ is expressed as

$$n^{-1}\mathbf{X}'\mathbf{Z} = \mathbf{B}'^{+}\mathbf{L}_1\Delta\mathbf{L}_1' = n^{-1}\mathbf{X}'\mathbf{XBL}_1\Delta^{-1}\mathbf{L}_1' \quad (8)$$

Proof. rank($\mathbf{XB}$) = p implies rank($\mathbf{B}$) = p and $\mathbf{BB}^{+} = \mathbf{I}_p$. Thus, (7) leads to the first equality in (8). Further, the pre-multiplying the left and right sides of (6) by $n^{-1}\mathbf{X}'$ and using $\mathbf{X}'\mathbf{K}_2\mathbf{L}_2' = \mathbf{B}'^{+}\mathbf{B}'\mathbf{X}'\mathbf{K}_2\mathbf{L}_2' = \mathbf{B}'^{+}\mathbf{L}_1'\Delta\mathbf{K}_1'\mathbf{K}_2\mathbf{L}_2' = \mathbf{O}$ (the zero matrix), we have the last equality in (8). □

Problem (5) is solved for $\mathbf{B} \equiv [\Lambda, \Psi] = [n^{-1}\mathbf{X}'\mathbf{F}, \text{diag}(n^{-1}\mathbf{X}'\mathbf{U})]$. It shows that the optimal $\mathbf{B}$ can be obtained using $n^{-1}\mathbf{X}'\mathbf{Z} \equiv n^{-1}\mathbf{X}'[\mathbf{F}, \mathbf{U}]$, which is given by (8). Thus, $\mathbf{B}$ can be uniquely determined in MDFA.

References

Adachi, K. (2012). Some contributions to data-fitting factor analysis with empirical comparisons to covariance-fitting factor analysis. *Journal of the Japanese Society of Computational Statistics*, **25**, 25-38.

Unkel, S. and Trendafilov, N. T. (2010). Simultaneous parameter estimation in exploratory factor analysis: An expository review. *International Statistical Review*, **78**, 363-382.

ten Berge, J. M. F. (1993). *Least squares optimization in multivariate analysis*. Leiden, The Netherlands: DSWO Press.