Sparsify Latent Factor Matrix by Householder Transformations

Xiaobai Sun

Abstract

In 1958 A. S. Householder (1904-1993) introduced the reflection transformation in his highly influential paper, *Unitary Triangularization of a Nonsymmetric Matrix*, published in the Journal of the ACM. He presented the reflection as a special case of nonsingular transformation matrices in the form of a rank-1 deviation from the identity matrix. In that same year, H. F. Kaiser (1927-1992) published the seminal paper *the Varimax Criterion for Analytic Rotations in Factor Analysis* in Psychometrika. Both papers have seen increasing citations in recent years, as will be demonstrated. This work introduces the use of Householder transformations for effective and efficient rotations and sparsification of latent factors. It has several advantages over the state-of-the-art factor rotation methods. This appears to be the first connection between these two lines of research. [1]

Analytic rotations are central to multiple factor analysis. Factor analysis is a statistical method to uncover one or more than one latent variables, a.k.a. factors, that explain or interpret the correlations among observable and observed variables. Latent variable analysis, originated from the pioneering work of C. Spearman in 1904 in psychology, is indispensible to modern exploratory analysis of data from various study fields, especially in social sciences and biomedical sciences. In multi-factor analysis, the relationship between the observable and latent variables is represented by a factor (loading) matrix. The concept of model simplification by factor rotations was conceived and developed between 1932 and 1938 by L. L Thurstone (1887-1955). Factor rotations are preceded by an initial factor extraction, which can be obtained manually based on expert knowledge or automatically via principal component analysis, maximal likelihood estimate, or other approaches. The factor axes are then re-oriented by orthogonal or oblique rotation transformations, to simplify (i.e. sparsify) the factor matrix pattern. The purpose is to identify salient relationships between the observed variables and the latent factors and to explain or interpret the correlation in observed phenomena. The term *simplification* here refers to reducing the complexity of observed variables in terms of the underlying factors. Thurstone's five simplification rules have been notably refined over time. Kaiser's varimax criterion and solution methods have a broad and lasting impact.

The factor rotation problem can be generally described as a factor matrix transformation governed by a constrained nonlinear optimization problem. An objective function specifies a simplification (or sparsification) criterion based on desired properties of the rotated factor matrix. All existing criteria, including Kaiser's criterion, are nonlinear functions with respect to the elements of the rotated factor matrix. Constraints include equations to ensure the orthogonality in orthogonal rotations or to preserve the variance per factor and avoid factor collapse in oblique rotations. Additional constraints may be imposed in confirmatory factor analysis to align with reference or target factor patterns.

Consider a particular case. Let $B_{p \times m}$ be the initial factor (loading) matrix with p variables and m factors, where $1 < m < p$. Let $L(U) = BU$ be the loading matrix rotated from B by an orthogonal transformation U. Kaiser's criterion can then be described as follows,

$$U_* = \arg \max_{U^{\mathrm{T}}U=I} \phi(L(U)) = \frac{e^{\mathrm{T}}(L.^4)e}{p} - \sum_{j=1:m} \left(\frac{L(:,j)^{\mathrm{T}}L(:,j)}{p} \right)^2, \qquad L(U) = BU, \qquad (1)$$

---

[1]This abstract is based on a manuscript not yet submitted anywhere to be considered for publication.

where e denotes the constant-1 vector[2]. The rotated factor matrix is $L(U_*)$.

Methods for computational solution of a factor rotation problem, such as (1), are inherently iterative due to the non-linearity of the objective criterion function. Kaiser's solution method involves iterative sweeps of plane rotations. Every sweep comprises $m(m-1)/2$ plane rotations across all pairwise factor axes. The single parameter for each plane rotation can be determined by a single equation derived from (1). Without being confined to the plane rotation sweeps, some methods take the alternative approach, which determines an $m \times m$ orthogonal matrix U, with $m(m-1)/2$ equations for the orthogonality. More specifically, one may use the Lagrange approach, with up to $m(m-1)/2$ multipliers, or deploy the gradient ascending method followed by a projection into the feasible solution space.

The specialized use of Householder transformations for factor matrix sparsification can effectively mitigate or eliminate certain issues present in existing factor rotation methods. For orthogonal factor rotations, a Householder reflection is used in place of a plane-rotation sweep as in Kaiser's method. In comparison to the alternative approach, this new approach implicitly decomposes a general orthogonal matrix U into orthogonal factors of a compact form. At each step, there are $(m-1)$ parameters to be determined, as opposed to just 1 at one extreme with Kaiser's method or $m(m-1)/2$ at the other extreme with the alternative approach. For any $m > 1$, there is only one Lagrange multiplier. The new factor rotation approach is simple in derivation as well as in implementation. It effectively eliminates the sequencing or scheduling problem within each sweep of plane rotations and resolves the projection issue encountered in gradient-based iteration methods. Additionally, numerical experiments, which will be presented, demonstrate that the new approach is more efficient. For oblique rotations, the use of an oblique Householder transformation is introduced, with similar benefits. Not restricted to Kaiser's criterion, the new method for sparsifying the factor matrix is compatible with and applicable to all factor rotation criteria commonly used in practice.

As a curious application, the simplification criteria and methods are also utilized to sparsify the base vectors for a multi-dimensional Householder reflection.

---

[2]The criterion will be explained in the presentation