A Note on Terse Coding of Kaiser's Varimax Rotation Using Complex Number Representation

Trevor Park

March 2003

1 Introduction

Kaiser (1958) both introduced the varimax criterion for orthogonal factor rotation and proposed a practical algorithm for its optimization, based on successive pairwise, or *planar*, rotations. (The idea was possibly inspired by similar methods introduced earlier for the quartimax criterion by, for instance, Neuhaus and Wrigley (1954).) That algorithm remains the standard method for varimax rotation in factor analysis (c.f. Harman 1976) and has been generalized to related rotation criteria (c.f. Clarkson and Jennrich 1988).

It is not widely appreciated that Kaiser's pairwise rotation solution for varimax can be much more compactly represented in terms of complex numbers. Such a connection allows much more terse coding of the algorithm in high-level computer languages capable of directly performing complex arithmetic. It also provides another verification of the formulas in the computational appendix of Kaiser (1958), where the derivation is only briefly and incompletely sketched due to the inordinate amount of algebraic manipulation involved.

2 Algorithmic Representation in Terms of Complex Numbers

In the context of factor analysis[1], consider a $d \times m$ loadings matrix $A = [a_{jk}]$ (where generally $d > m$). The varimax criterion value for this matrix is given by

$$V(A) \;=\; \sum_k \left(\tfrac{1}{d} \sum_j a_{jk}^4 - \left(\tfrac{1}{d} \sum_j a_{jk}^2 \right)^2 \right). \tag{1}$$

(Kaiser (1958) refers to this as "raw" varimax, but it is the version that has become most popular.) Verbally, this is simply the sum of the columnwise variances of the squared elements of A.

An orthogonal rotation of a loadings matrix A is any matrix of the form AR, where R is an $m \times m$ orthogonal matrix. A varimax-rotated version of a loadings matrix A is any orthogonal rotation of A that maximizes the varimax criterion among all orthogonal rotations of A. Such a matrix will generally have fewer large-magnitude loadings in its columns, thereby making the columns more easily interpretable in terms of the variables corresponding to the rows.

The pairwise (or planar) rotation algorithm for optimizing the varimax criterion of AR over all orthogonal matrices R proceeds as specified on the following page. It bears a slight resemblance to Jacobi-style methods for symmetric eigenvalue problems in that it produces a sequence of iterates converging to a final solution, each iterate obtained by applying a Givens rotation to the previous iterate (c.f. Golub and Van Loan 1996, Sec. 8.4). Little is known for certain about the convergence

[1]This article uses some terminology associated with factor analysis, but it should still be accessible to someone without such a background. Rotation criteria like varimax are sometimes used in other types of multivariate exploratory analysis, such as principal component analysis.

Set A to the original loadings matrix
Repeat
For each unordered pair of column indices (k, l), visited in some arbitrary order,
Find a 2×2 orthogonal matrix R^* that achieves $\max_R V([\mathbf{a}_k \, \mathbf{a}_l] R)$,
where $\mathbf{a}_k$ and $\mathbf{a}_l$ are columns k and l of A. (See text for details.)
Replace these columns of A with the columns of $[\mathbf{a}_k \, \mathbf{a}_l] R^*$.
Until convergence

behavior of the algorithm, but in practice it usually converges to a local maximizer of the varimax criterion.

The key step in the algorithm is solving the maximization problem $\max_R V([\mathbf{a}_k \, \mathbf{a}_l] R)$, over all 2×2 orthogonal matrices R, which is really a simple $m = 2$ case of the varimax rotation problem. Kaiser (1958) showed that a simple analytical solution exists in the $m = 2$ case, thus allowing the algorithm to operate efficiently. Using the notation

$$t = 2\left(d \sum_j (a_{jk}^2 - a_{jl}^2)(2a_{jk}a_{jl}) - \sum_j (a_{jk}^2 - a_{jl}^2) \sum_j (2a_{jk}a_{jl})\right)$$

$$b = d \sum_j \left((a_{jk}^2 - a_{jl}^2)^2 - (2a_{jk}a_{jl})^2\right) - \left(\sum_j (a_{jk}^2 - a_{jl}^2)\right)^2 - \left(\sum_j (2a_{jk}a_{jl})\right)^2,$$

the optimizing rotation angle, as given by Kaiser, is

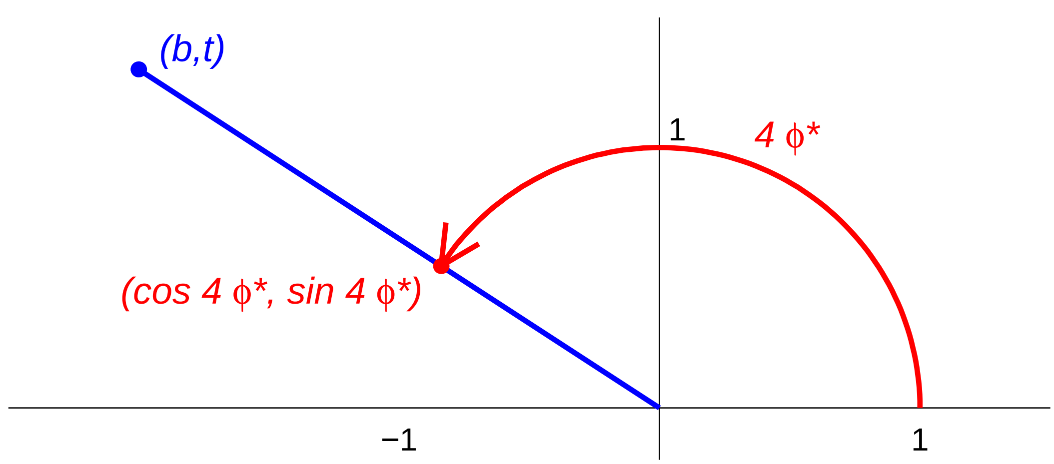
$$\phi^* = \tfrac{1}{4} \arctan(t, b),$$

where $\arctan(y, x)$ is the *four-quadrant* inverse tangent, i.e. the polar-coordinate angle of the point (x, y) in the Cartesian plane (see Figure 1). (See also Clarkson and Jennrich (1988).) The optimal rotation matrix R^* is then given by

$$R^* = \begin{bmatrix} \cos(\phi^*) & -\sin(\phi^*) \\ \sin(\phi^*) & \cos(\phi^*) \end{bmatrix}. \quad (2)$$

Explicit expressions for the elements of R^* directly in terms of t and b that avoid transcendental functions can be derived by using standard trigonometric identities (Nevels 1986; Harman 1976, p. 294).

Figure 1: Schematic illustrating the relationship between t, b, and ϕ^*.



If the planar varimax rotation problem is embedded in the complex plane, a relatively simple and compact alternative expression for ϕ^* directly in terms of the elements of A is

$$\phi^* \;=\; \tfrac{1}{4}\angle\Big(\tfrac{1}{d}\sum_j (a_{jk}+ia_{jl})^4 \;-\; \Big(\tfrac{1}{d}\sum_j (a_{jk}+ia_{jl})^2\Big)^2\Big), \tag{3}$$

where $\angle$ stands for the angle in the complex plane. The derivation of this expression is given in Section 3. (Note that the complex number in expression (3) is apparently the inner expression of the varimax criterion (1) applied to the complex number $a_{jk}+ia_{jl}$.) Once ϕ^* is computed, R^* can be obtained from (2). It may be possible to compute the elements of R^* without trigonometric functions using the complex number approach, but the author has not pursued this topic, as evaluations of trigonometric functions are relatively fast on modern computers.

It should be easy to compactly code expression (3) in high-level numerical languages that can perform complex arithmetic. For instance, in the popular MATLAB[2] language, expression (3) could be coded as

```
phi_max = angle(sum(complex(A(:,k),A(:,l)).^4)/d  ...
                - (sum(complex(A(:,k),A(:,l)).^2)/d)^2) / 4;
```

In S-PLUS[3], it could be coded as

```
phi.max <- Arg(sum(complex(re=A[,k],im=A[,l])^4)/d
               - (sum(complex(re=A[,k],im=A[,l])^2)/d)^2) / 4
```

Of course, the advantage of this approach is in terse coding, not computational efficiency. For low-level languages, the standard algorithms are preferable.

3 Derivation of Planar Varimax Solution Using Complex Numbers

Consider the problem of optimal varimax rotation of the $d \times 2$ loadings matrix

$$A \;=\; \begin{bmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_d & y_d \end{bmatrix}.$$

For this matrix A, denote the complex number in expression (3) as

$$C_A \;=\; \tfrac{1}{d}\sum_j (x_j+iy_j)^4 \;-\; \Big(\tfrac{1}{d}\sum_j (x_j+iy_j)^2\Big)^2.$$

The real part of this number can be expanded as follows:

$$\begin{aligned}\mathrm{Re}(C_A) \;&= 4\,V(A) -\; 3\,\tfrac{1}{d}\sum_j (x_j^2+y_j^2)^2 \\ &\quad +\; \Big(\tfrac{1}{d}\sum_j (x_j^2+y_j^2)\Big)^2 \;+\; 2\,\tfrac{1}{d^2}\sum_j\sum_{j\prime}(x_jx_{j\prime}+y_jy_{j\prime})^2.\end{aligned} \tag{4}$$

The remarkable aspect of this expansion is that all terms except $4\,V(A)$ are invariant to rotations AR of A, where R is 2×2 orthogonal. To see this, note that the remaining terms involve only

[2]MATLAB is a registered trademark of The MathWorks, Inc.
[3]S-PLUS is a registered trademark of Insightful Corporation.

expressions of the form $x_j^2 + y_j^2$ and $x_j x_{j'} + y_j y_{j'}$, which are the elements of the matrix $AA^\mathsf{T} = (AR)(AR)^\mathsf{T}$ for any orthogonal R. Thus, for purposes of 2×2 orthogonal rotation, (4) is equivalent to varimax.

Now, define

$$R(\phi) \;=\; \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}, \tag{5}$$

and let

$$\tilde{A}(\phi) \;=\; \begin{bmatrix} \tilde{x}_1(\phi) & \tilde{y}_1(\phi) \\ \vdots & \vdots \\ \tilde{x}_d(\phi) & \tilde{y}_d(\phi) \end{bmatrix} \;=\; AR(\phi). \tag{6}$$

The varimax rotation problem for A is then the problem of finding ϕ^* that maximizes $V(\tilde{A}(\phi))$. From the foregoing, this is equivalent to maximizing

$$\mathrm{Re}\left(\tfrac{1}{d} \sum_j (\tilde{x}_j(\phi) + i\tilde{y}_j(\phi))^4 \;-\; \left(\tfrac{1}{d} \sum_j (\tilde{x}_j(\phi) + i\tilde{y}_j(\phi))^2 \right)^2 \right) \tag{7}$$

in ϕ. But (5) and (6) imply the relationship

$$\tilde{x}_j(\phi) + i\tilde{y}_j(\phi) \;=\; (x_j + iy_j)\, e^{-i\phi}$$

and so (7) becomes

$$\mathrm{Re}\left(\tfrac{1}{d} \sum_j ((x_j + iy_j)\, e^{-i\phi})^4 \;-\; \left(\tfrac{1}{d} \sum_j ((x_j + iy_j)\, e^{-i\phi})^2 \right)^2 \right) \;=\; \mathrm{Re}(e^{-i4\phi}\, C_A). \tag{8}$$

Taking the first derivative of (8) in ϕ and setting it equal to zero gives the first-order necessary condition for the maximizer ϕ^*:

$$\mathrm{Re}(-i4e^{-i4\phi^*}\, C_A) \;=\; 4\,\mathrm{Im}(e^{-i4\phi^*}\, C_A) \;=\; 0,$$

that is,

$$C_A \;=\; re^{i4\phi^*}$$

for some real r. If r is positive, it will immediately follow that

$$\phi^* \;=\; \tfrac{1}{4}\,\angle(C_A),$$

thus verifying (3). To show that r is positive, take the second derivative of (8) in ϕ and use the second-order condition for the maximizer ϕ^*:

$$\mathrm{Re}((-i4)^2 e^{-i4\phi^*}\, C_A) \;=\; -16\,\mathrm{Re}(e^{-i4\phi^*}\, C_A) \;=\; -16r \;<\; 0.$$

In passing, note that

$$\frac{d}{d\phi} V(\tilde{A}(\phi)) \;=\; \mathrm{Im}(e^{-i4\phi}\, C_A),$$

as follows from considerations involving formulas (4) and (8). This is the essential relationship behind the foregoing derivation.

References

Clarkson, D. B., and Jennrich, R. I. (1988), "Quartic Rotation Criteria and Algorithms," *Psychometrika*, 53, 251–259.

Golub, G. H., and Van Loan, C. F. (1996), *Matrix Computations* (3rd ed.), Baltimore, MD: Johns Hopkins University Press.

Harman, H. H. (1976), *Modern Factor Analysis* (rev. 3rd ed.), Chicago: University of Chicago Press.

Kaiser, H. F. (1958), "The Varimax Criterion for Analytic Rotation in Factor Analysis," *Psychometrika*, 23, 187–200.

Neuhaus, J. O., and Wrigley, C. (1954), "The Quartimax Method: An Analytic Approach to Orthogonal Simple Structure," *British Journal of Statistical Psychology*, 7, 81–91.

Nevels, K. (1986), "A Direct Solution for Pairwise Rotations in Kaiser's Varimax Method," *Psychometrika*, 51, 327–329.