ORTHOGONAL ROTATION TO A
PARTIALLY SPECIFIED TARGET*

Franklin T. Luk

TR 81-470
September 1981

Department of Computer Science
Cornell University
Ithaca, New York 14853

*This research was supported in part by the U.S. Army Research Office under grant DAAG 29-79-C0124

Orthogonal Rotation to a Partially Specified Target

Franklin T. Luk*

Department of Computer Science

Cornell University

Ithaca, New York 14853

Abstract

This paper addresses the problem of finding an orthogonal transformation of an arbitrary factor solution that would lead to a least squares fit of a partially specified target matrix. An iterative computing procedure is presented.

Key words: orthogonal rotations, factor analysis, least squares, partially specified target, procrustes problem.

* This research was supported in part by the U.S. Army Research Office under grant DAAG 29-79-C0124.

1. Introduction

We consider the problem of orthogonally rotating a given factor matrix to best fit a partially prescribed factor pattern in the least squares sense. Lawley and Maxwell [4] present an algorithm for the special case where all the target elements are zero. Browne [1] considers the general case and proposes the application of plane rotations to compute an approximate solution. The rotations are determined by Newton's method which requires very good initial guesses. In this paper we show that these planar rotations are expressible as solutions to linear least squares problems with an equality constraint. Effective numerical techniques are therefore applicable, and a procedure based on the singular value decomposition (cf. Golub [3]) will be presented.

2. The Algorithm

Our problem is frequently referred to as a procrustes problem, a term borrowed from Greek mythology in which a highwayman named Procrustes is supposed to have made all his victims fit his bed, cruelly stretching the short ones and lopping off the tall ones. Mathematically, our problem is to determine an $m \times m$ orthogonal matrix Q that will transform a given $n \times m$ factor matrix A to best fit a partially prescribed $n \times m$ target factor pattern B in the least squares sense. Let $I(j)$ represent the set of row indices corresponding to specified elements of column j of B. If $F = AQ$, the criterion to be minimized is

$$(2.1) \quad g \equiv \sum_{j=1}^{m} \{ \sum_{i \in I(j)} (f_{ij} - b_{ij})^2 \},$$

where $F = (f_{ij})$ and $B = (b_{ij})$.

Following Browne [1] we shall approximate Q by a product of plane rotations. Designate by A^c the current rotated factor matrix at an intermediate stage, and let $A^c = (a^c_{ij}) = (\mathbf{a}^c_1, \mathbf{a}^c_2, \ldots, \mathbf{a}^c_m)$. We now consider a rotation over the plane defined by the i-th and j-th factors:

$$(2.2) \qquad (\mathbf{a}^c_i, \mathbf{a}^c_j)\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} .$$

Let i(1),i(2),...,i(I(i)) and j(1),j(2),...,j(I(j)) denote the row indices of the specified elements of columns i and j, respectively, of the target B. Our goal is to determine a rotational angle θ to minimize the euclidean length of the vector

$$(2.3) \qquad \begin{bmatrix} a^c_{i(1),i} & a^c_{i(1),j} \\ a^c_{i(2),i} & a^c_{i(2),j} \\ \vdots & \vdots \\ a^c_{i(I(i)),i} & a^c_{i(I(i)),j} \\ a^c_{j(1),j} & -a^c_{j(1),i} \\ a^c_{j(2),j} & -a^c_{j(2),i} \\ \vdots & \vdots \\ a^c_{j(I(j)),j} & -a^c_{j(I(j)),i} \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} - \begin{bmatrix} b_{i(1),i} \\ b_{i(2),i} \\ \vdots \\ b_{i(I(i)),i} \\ b_{j(1),j} \\ b_{j(2),j} \\ \vdots \\ b_{j(I(j)),j} \end{bmatrix} .$$

In other words, we want to solve the constrained least squares problem:

$$(2.4) \qquad \|C\mathbf{x} - \mathbf{d}\|_2 = \text{minimum},$$

subject to

$$(2.5) \qquad \|\mathbf{x}\|_2 = 1.$$

Effective algorithms for solving this problem are well known. We shall describe one based on the singular value decomposition (cf. Golub [3]).

The singular value decomposition of an $n \times 2$ matrix C is readily computable. First, two Householder transformations, say H_1 and H_2, can be applied to reduce the matrix to upper triangular form, i.e.

$$(2.6) \qquad H_2H_1C = \begin{bmatrix} p & q \\ 0 & r \\ 0 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 0 \end{bmatrix} \equiv \begin{bmatrix} R \\ \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} .$$

We are interested in orthogonalizing the two rows of R. As this problem is trivial if $pqr = 0$, we assume that all three elements are nonzero. As in Luk [5] we define

$$(2.7) \qquad \alpha \equiv 2qr, \ \beta \equiv p^2+q^2-r^2, \ \gamma \equiv (\alpha^2+\beta^2)^{1/2}.$$

If $\beta > 0$, we compute

$$(2.8) \qquad c = \left(\frac{\gamma+\beta}{2\gamma}\right)^{1/2}, \ s = \frac{\alpha}{2\gamma c}.$$

else we compute

$$(2.9) \qquad s = \left(\frac{\gamma-\beta}{2\gamma}\right)^{1/2}, \ c = \frac{\alpha}{2\gamma s}.$$

Defining

$$(2.10) \qquad Q \equiv \begin{pmatrix} c & -s \\ s & c \end{pmatrix},$$

we have

$$(2.11) \qquad Q^TR = \begin{pmatrix} cp & cq+sr \\ -sp & -sq+cr \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix}$$

$$\equiv \Sigma V^T,$$

where

$$\sigma_1 = [c^2p^2 + (cq+sr)^2]^{1/2},$$

$$\sigma_2 = [s^2p^2 + (-sq+cr)^2]^{1/2},$$

$$v_{11} = cp/\sigma_1,$$

$$v_{12} = -sp/\sigma_2,$$

$$v_{21} = -v_{12},$$

and $\quad v_{22} = v_{11}.$

The singular value decomposition of C is then given by

(2.12) $\quad C = U\Sigma V^T,$

where

$$U = \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix} H_1H_2.$$

We shall use the notations

(2.13) $\quad U = (\mathbf{u}_1, \mathbf{u}_2)$

and

(2.14) $\quad V = (\mathbf{v}_2, \mathbf{v}_2).$

Let us now consider the constrained least squares problem (2.4) and

(2.5). The case of a null $\mathbf{d}$ has the solution $\mathbf{x} = \mathbf{v}_2$. Henceforth we shall assume that $\mathbf{d} \neq \mathbf{0}$. For a zero matrix C, we choose the solution $\mathbf{x} = (1, 0)^T$, i.e. the null rotation. If C is of rank one, then

$$(2.15) \qquad C^+\mathbf{d} = (\mathbf{u}_1^T\mathbf{d}/\sigma_1)\mathbf{v}_1 \equiv \xi\,\mathbf{v}_1 \;.$$

For $|\xi| < 1$, we let

$$(2.16) \qquad \mathbf{x} = \xi\mathbf{v}_1 + (1-\xi^2)^{1/2}\,\mathbf{v}_2,$$

else we let

$$(2.17) \qquad \mathbf{x} = \text{sign}\,(\xi)\;\mathbf{v}_1.$$

If C is of full rank, we consider the matrix equation

$$(2.18) \qquad (C^TC+\lambda I)\mathbf{x} = C^T\mathbf{d},$$

where λ is a Lagrange multiplier. The proper value for λ is a solution to the nonlinear equation

$$(2.19) \qquad \det[(\Sigma^T\Sigma + \lambda I)^2 - \mathbf{e}\mathbf{e}^T] = 0,$$

where

$$(2.20) \qquad \mathbf{e} = \Sigma^TU^T\mathbf{d} \equiv \binom{e_1}{e_2}.$$

Although equation (2.19) is quartic, it can be shown (cf. Forsythe and Golub [2]) that we need the largest real root λ_1. Furthermore, if

$$||C^+\mathbf{d}||_2 > 1,$$

then λ_1 is the unique root in the open interval $(0, ||\mathbf{e}||_2)$, otherwise, except for a singular case, λ_1 is the unique root in the semi-open interval $(-\sigma_2^2, 0]$. We may use linear interpolation to compute λ_1. The solution vector is given by

$$(2.21) \qquad \mathbf{x} = V(\Sigma^T\Sigma + \lambda_1 I)^{-1}\mathbf{e}.$$

The singularity occurs when $\sigma_1 > \sigma_2, e_2 = 0$ and $|e_1| \le (\sigma_1^2 - \sigma_2^2)$. We then have $\lambda_1 = -\sigma_2^2$ and

$$(2.22) \qquad \mathbf{x} = \left(\frac{\mathbf{e}}{\sigma_1^2 - \sigma_2^2}\right)\mathbf{v}_1 + \xi\mathbf{v}_2,$$

where ξ is chosen such that $||\mathbf{x}||_2 = 1$.

The plane rotations are applied to the factor matrix in a systematic order. Transformations are made on factors 1 with 2, 1 with 3, ...,1 with m, 2 with 3, ..., (m-1) with m to constitute a cycle. The cycle is repeated until either there is one consisting of all zero rotations or the criterion g of (2.1) fails to decrease. The criterion should decrease after each non-trivial rotation, and as it is bounded below by zero, the termination of our algorithm is guaranteed.

We should mention that Browne [1] also contains suggestions on when to interchange or reflect the factors.

3. Numerical Example

We select the example in Browne [1]. The given factor matrix is

$$A = \begin{bmatrix} .664 & .322 & -.075 \\ .688 & .248 & .192 \\ .492 & .304 & .224 \\ .837 & -.291 & .037 \\ .705 & -.314 & .155 \\ .820 & -.377 & -.104 \\ .661 & .397 & .077 \\ .457 & .294 & .488 \\ .765 & .428 & .009 \end{bmatrix}$$

and the partially prescribed matrix is

$$B = \begin{bmatrix} x & 0 & x \\ x & 0 & 0 \\ x & 0 & 0 \\ x & x & x \\ x & x & 0 \\ x & x & x \\ .7 & x & x \\ 0 & x & x \\ .7 & x & x \end{bmatrix} .$$

We regard a rotation as null if $\cos\theta \geq .9999995$. The sixth cycle consists wholly of null rotations and we obtain the final rotated matrix

$$AQ = \begin{bmatrix} .611 & -.023 & -.420 \\ .734 & -.055 & -.173 \\ .607 & .086 & -.093 \\ .608 & -.622 & -.176 \\ .549 & -.564 & -.017 \\ .499 & -.713 & -.261 \\ .705 & .069 & -.314 \\ .235 & -.024 & -.691 \\ .768 & .039 & -.422 \end{bmatrix} .$$

The criterion decreases as follows:

Cycle	Criterion
0	.583264
1	.182168
3	.113184
3	.110034
4	.109858
5	.109852
6	.109852

Acknowledgements. The author is grateful to C.F. Van Loan for many valuable suggestions.

References

[1] M.W. Browne, "Orthogonal rotation to a partially specified target," *Br. J. math. statist. Psychol.* *25* (1972), 115-120.

[2] G.E. Forsythe and G.H. Golub, "On the stationary values of a second-degree polynomial on the unit sphere," *J. Soc. Indust. Appl. Math.* *13* (1965), 1050-1068.

[3] G.H. Golub, "Some modified matrix eigenvalue problems," *SIAM Review* *15* (1973), 318-334.

[4] D.N. Lawley and A.E. Maxwell, "Factor transformation methods," *Br. J. statist. Psychol.* *17* (1964), 97-103.

[5] F.T. Luk, "Computing the singular-value decomposition on the ILLIAC IV," *ACM Trans. Math. Softw.* *6* (1980), 524-539.