

A SOLUTION TO THE WEIGHTED PROCRUSTES PROBLEM
IN WHICH THE TRANSFORMATION IS IN
AGREEMENT WITH THE LOSS FUNCTION

ROBERT W. LISSITZ

UNIVERSITY OF GEORGIA

PETER H. SCHÖNEMANN

PURDUE UNIVERSITY

JAMES C. LINGOES

UNIVERSITY OF MICHIGAN

This paper provides a generalization of the Procrustes problem in which the errors are weighted from the right, or the left, or both. The solution is achieved by having the orthogonality constraint on the transformation be in agreement with the norm of the least squares criterion. This general principle is discussed and illustrated by the mathematics of the weighted orthogonal Procrustes problem.

Key words: rotation, matching.

In Schönemann [1966], a general solution was given for least squares problems of the type

$$(1) \quad B = AT + E,$$

where A and B are two given matrices of order $p \times m$ ($m < p$). The objective is to find an orthogonal $m \times m$ matrix T which minimizes the sum of squared residuals; i.e.,

$$(2) \quad \min_r \text{tr } E'E \text{ so that } T'T = TT' = I_m.$$

The solution for this "orthogonal Procrustes" problem is

$$(3) \quad T = VW', \text{ where}$$

$$(4) \quad C = A'B = VDW';$$

i.e., V and W are the vectors (which are, if necessary, orthogonalized) in the Eckart and Young [1936] decomposition of $C = A'B$. Technical problems associated with this solution are discussed in Schönemann [1966] and Schönemann, Bock, and Tucker [Note 2]. Schönemann and M. Carroll [1970] have extended this solution to include translations and central

Requests for reprints should be sent to Robert W. Lissitz, Department of Psychology, University of Georgia, Athens, Georgia 30602.

dilations, while Lingoes and Schönemann [1974] have developed a symmetric and dimensionless measure S to assess the fit.

Some applications call for an extension in a different direction. Instead of working with the uniformly weighted loss function $\text{tr } E'E$ in (2), one may wish to weight either the rows or columns (of E) or both differentially in the loss function [Lingoes, in press]. Technically, this means that one replaces the Euclidean distance function $\text{tr } E'E$ in (2) by some weighted ("elliptical") distance function, e.g.

$$(5) \quad \text{tr } E'D_1^2E,$$

or

$$(6) \quad \text{tr } ED_2^2E',$$

where D_1 is any matrix and D_2 is any positive definite matrix. For most applications, these would probably be diagonal matrices. There is no particular problem with the first case (5), which may be thought of as a change in the A and B matrices. If we write

$$(7) \quad B = AT' + E,$$

and

$$(8) \quad \min_T \text{tr } E'D_1^2E \quad \text{so that } T'T = TT' = I,$$

we have

$$(9) \quad D_1E = D_1B - D_1AT = B^* - A^*T = E^*, \\ A^* = D_1A,$$

and

$$B^* = D_1B,$$

where $T = V^*W^{**}$ as in (3) and $C^* = A^*B^* = V^*DW^{**}$ and $T'T = TT' = I$ as in (4). However, in the second case (6), the resulting model becomes unwieldy. Again,

$$(10) \quad B = AT' + E,$$

but now we seek

$$(11) \quad \min_T \text{tr } ED_2^2E', \quad \text{so that } T'T = TT' = I.$$

The function to be minimized is

$$(12) \quad f = \text{tr } ED_2^2E' - 2 \text{tr } L(T'T - I).$$

Differentiating (see, e.g., Schönemann [Note 1]) with respect to T , we find

$$(13) \quad 2 \partial f / \partial T = A'B - A'ATD - TQ = \phi,$$

where L is an unknown matrix of Lagrange multipliers so that

$$(14) \quad Q = L + L' = \text{symmetric}.$$

In this case there does not seem to be any simple solution for T and Q since the second

term in (13) cannot be eliminated by appeal to the symmetry of Q , which was the key for the solution of the unweighted problem given by (1) and (2).

A tractable solution for such weighted least squares problems can be obtained if one sees to it that the scalar product function, $f(x, y)$, for the orthogonality constraint on T agrees with the norm, $n(x)$, for defining the least squares criterion so that $f(x, x) = n^2(x)$ for all vectors x . To illustrate this principle, we solve the problem of (6), where

$$(15) \quad ED_2 = BD_2 - ATD_2 = E^*,$$

and

$$(16) \quad \min_{T^*} \text{tr } ED_2^2E' = \min_{T^*} \text{tr } E^*E^{*'}.$$

A minimization of $\text{tr } E^*E^{*'} = \text{tr } ED_2^2E'$ (under choice of T) can be viewed as a minimization of the squared lengths of the row vectors E_i' in E with respect to the weighted norm $E_i'D_{2i}^2E_i$. To render T "orthogonal" with respect to the associated scalar product $E_i'D_{2i}^2E_i$ for the rows E_i' in E , one would have to impose

$$(17) \quad TD_2^2T' = T^*T^{*'} = T^{*'}T^* = I,$$

where $T^* = TD_2$ and where the commutativity is required for orthogonality. If one now differentiates

$$(18) \quad f^* = \text{tr } E^*E^{*'} - 2 \text{tr } L(T^{*'}T^* - I),$$

with respect to T^* and sets the derivative equal to zero, one obtains

$$(19) \quad T^{*'}A'BD_2 = L + L' = \text{symmetric},$$

analogous to the unweighted case. Hence

$$(20) \quad T^* = VW', \quad \text{or} \quad T = T^*D_2^{-1},$$

is the solution, where V and W contain the vectors of the Eckart-Young decomposition of

$$(21) \quad C = A'BD_2 = VDW'.$$

This solution for the weighted least squares problem is of the same form as the solution for the unweighted problem because T^* , not T , is a rotation in a Euclidean space where $\text{tr } E^{*'}E^* = \text{tr } E^*E^{*'}$ is a sum of squared Euclidean distances. To achieve this, we had to amend the side condition as in (17) so that the scalar product which defines the orthogonality condition corresponds with the norm which defines the least squares criterion. The only difference between the solutions is the matrix C in (4) and (21). If the researcher were interested in weighting by D_1 and D_2 simultaneously, then the A^* and B^* as in (9) would be the input to the solution of (2) instead of the A and B as given. The program MFIT in the $G - L$ series [Lingoes, 1973] has been modified to take into account either, both, or neither weighting system.

REFERENCE NOTES

1. Schönemann, P. H. *On the formal differentiation of traces and determinants* (Res. Memo No. 27). Chapel Hill: University of North Carolina Psychometric Laboratory, 1965.
2. Schönemann, P. H., Bock, R. D., and Tucker, L. R. *Some notes on a theorem by Eckart and Young* (Res. Memo. No. 25). Chapel Hill: University of North Carolina Psychometric Laboratory, 1965.

REFERENCES

- Eckart, C. and Young, G. The approximation of one matrix by another of lower rank. *Psychometrika*, 1936, 1, 211-218.
- Green, B. F. The orthogonal approximation of an oblique simple structure in factor analysis. *Psychometrika*, 1952, 17, 429-440.
- Kristof, W. Die beste orthogonale transformation zur gegenseitigen Überführung zweier Faktorenmatrizen. *Diagnostica*, 1964, 10, 87-90.
- Lingoes, J. C. *The Guttman-Lingoes nonmetric program series*. Ann Arbor, Michigan: Mathesis Press, 1973.
- Lingoes, J. C. and Schönemann, P. H. Alternative measures of fit for the Schönemann-Carroll matrix fitting algorithm. *Psychometrika*, 1974, 39, 423-427.
- Lingoes, J. C. A neighborhood preserving transformation for fitting configurations. In J. C. Lingoes, L. Guttman, and E. Roskam (Eds.), *Geometric representations of relational data: With social science applications*. Ann Arbor, Michigan: Mathesis Press, in press.
- Schönemann, P. H. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 1966, 31, 1-10.
- Schönemann, P. H. and Carroll, R. M. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, 1970, 35, 245-255.

Manuscript received 12/11/75

Final version received 5/18/76