SORT 28 (1) January-June 2004, 27-36

On best affine unbiased covariance-preserving prediction of factor scores

Heinz Neudecker*

University of Amsterdam, The Netherlands

Abstract

This paper gives a generalization of results presented by ten Berge, Krijnen, Wansbeek & Shapiro. They examined procedures and results as proposed by Anderson & Rubin, McDonald, Green and Krijnen, Wansbeek & ten Berge. We shall consider the same matter, under weaker rank assumptions. We allow some moments, namely the variance Ω of the observable scores vector and that of the unique factors, Ψ, to be singular. We require $T'\Psi T > 0$, where $T\Lambda T'$ is a Schur decomposition of Ω. As usual the variance of the common factors, Φ, and the loadings matrix A will have full column rank.

MSC: 62H25, 15A24

Keywords: Factor analysis, factor scores, covariance-preserving, Kristof-type theorem

1 Introduction

We consider the factor model $y = \mu_y + Af + \varepsilon$, where y is a $p \times 1$ vector of observable random variables called «scores», f is an $m \times 1$ vector of non-observable random variables called «common factors», A is a $p \times m$ matrix of full column rank whose elements are called «factor loadings» and ε is a $p \times 1$ vector of non-observable random variables called «unique factors». The usual moment definitions and assumptions are

$$E(\varepsilon) = 0, \quad E(f) = 0, \quad E(y) = \mu_y, \quad D(\varepsilon) = \Psi, \quad D(f) = \Phi, \quad C(f, \varepsilon) = 0.$$

* *Address for correspondence:* Oosterstraat, 13. 1741 GH Schagen. The Netherlands. Postal address: NL 1741 GH 13. E-mail: H.Neudecker@uva.nl

Received: December 2001
Accepted: September 2003

This yields the moment structure

$$\Omega = A\Phi A' + \Psi,$$

where $\Omega = D(y)$ and Ψ can be singular, Φ and A have full column rank.
Notice that

$$\mathcal{M}(A) \subset \mathcal{M}(\Omega). \tag{1.1}$$

The following *additional* assumption is made:

$$T'\Psi T > 0.$$

It is inspired by the Schur decomposition $\Omega = T\Lambda T'$, with $T'T = I_r$ and diagonal $\Lambda > 0$. Obviously $p \geqslant r > m$.

In two recent publications Krijnen, Wansbeek & ten Berge (1996) and ten Berge, Krijnen, Wansbeek & Shapiro (1999) studied the problem of best linear prediction of f given y, subject to the constraint $E\hat{f}\hat{f}' = Eff'$, where $\hat{f} = B'y$ is their predictor function. Vectors f and y have a simultaneous distribution. The two expectations are taken with respect to this distribution.

The constraint $E\hat{f}\hat{f}' = Eff'$ is mistakenly referred to as «correlation-preserving». We shall call it «covariance-preserving», although at face value only the RHS expression is a variance matrix. We shall use an affine predictor function $\hat{f} = a + B'y$. It will be shown that $a + B'\mu_y = 0$. Hence the predictor function will become $\hat{f} = B'(y - \mu_y)$ which is linear and unbiased. Consequently the LHS expression will become a variance matrix.

In their article ten Berge *et al.* (1999) examine three prediction procedures, due to McDonald (1981) —who generalized a procedure proposed by Anderson & Rubin (1956)—, Green (1969) and Krijnen *et al.* (1996), respectively.

We shall consider the same three procedures. The second and third are based on the mean-squared-error matrix $M = E\left(\hat{f} - f\right)\left(\hat{f} - f\right)'$. Where Green minimizes its trace, $\operatorname{tr} M$, Krijnen *et al.* minimize its determinant, $|M|$. McDonald uses a different though related criterion $\operatorname{tr}\Psi^{-1}E\left(y - \mu_y - A\hat{f}\right)\left(y - \mu_y - A\hat{f}\right)'$ which he minimizes. Note that these authors assume $\Psi > 0$, hence $\Omega > 0$. ten Berge *et al.* conclude that McDonald's and Krijnen *et al.*'s solutions for B coincide.

In the present paper we shall again consider the above-mentioned procedures, under weaker rank assumptions. We shall show that the MSE matrix M is positive definite. Minimization of the trace and the determinant of M yields immediately $a + B'\mu_y = 0$. Minimization of McDonald's criterion function yields the same result. As mathematical methods we use a Kristof-type theorem and a matrix inequality developed by Zhang (1999). Finally we show that 1) $\hat{f}_G$, the Green predictor and $\hat{f}_K$, the Krijnen *et al.* predictor coincide when Φ and $A'\Omega^+A$ commute, 2) $\hat{f}_M$, the McDonald predictor and $\hat{f}_K$ coincide when Ψ and $A\Phi A'$ commute.

2 A Kristof-type theorem

Two of the three criterion functions can be seen to belong to the class $\operatorname{tr} P'X$, where P and X have dimension $p \times m$. The constant matrix P has rank q. The variable matrix X satisfies the constraint $X'X = I_m$. The aim is to maximize $\operatorname{tr} P'X$ subject to $X'X = I_m$. Define then the Lagrangean function

$$\varphi(X) = \operatorname{tr} P'X - \tfrac{1}{2} \operatorname{tr} L\,(X'X - I_m)\,,$$

where L is a *symmetric* matrix of multipliers. Symmetry of L is vital. It is justified, of course, by the symmetry of the constraint.

The differential of the function, namely

$$d\varphi = \operatorname{tr} P'dX - \operatorname{tr} LX'dX = \operatorname{tr}\,(P - XL)'\,dX$$

has to be zero. This yields the equations

$$P = XL \tag{2.1}$$

$$X'X = I_m \tag{2.2}$$

From these we obtain

$$P'P = L^2 \tag{2.3}$$

$$P = X\,(P'P)^{\frac{1}{2}} \tag{2.4}$$

Which square root will be selected is still undecided. Consider equation (2.4). As

$$P\,(P'P)^{+\frac{1}{2}}\,(P'P)^{\frac{1}{2}} = P$$

it is consistent. The symbol «+» denotes the Moore-Penrose inverse. The symbols «+» and «$\frac{1}{2}$» are interchangeable in $(P'P)^{+\frac{1}{2}}$. The general solution of (2.4) is

$$X_\circ = P\,(P'P)^{+\frac{1}{2}} + Q - Q\,(P'P)^{\frac{1}{2}}\,(P'P)^{+\frac{1}{2}}\,, \quad Q \text{ arbitrary} \tag{2.5}$$

When we use the singular-value decomposition $P = F_1\Gamma_1^{\frac{1}{2}}G_1'$, with $F_1'F_1 = G_1'G_1 = I_q$ and (diagonal) $\Gamma_1^{\frac{1}{2}} > 0$, we can write the solution as

$$X_\circ = F_1G_1' + Q\,(I_m - G_1G_1') \tag{2.6}$$

It follows from (2.5) that

$$\operatorname{tr} P'X_\circ = \operatorname{tr}\,(P'P)^{\frac{1}{2}}\,. \tag{2.7}$$

As we look for a *maximum*, we have to take the *positive* definite square root $(P'P)^{\frac{1}{2}}$. The solution $X_\circ$ is *not* unique, unless $q = m$. In that case it can be written as

$$X_\circ = P\,(P'P)^{-\frac{1}{2}} = F_1G_1' \tag{2.8}$$

For the connaisseurs we shall examine the second differential

$$d^2\varphi = -\mathrm{tr}(dX)L(dX)' \tag{2.9}$$

When this expression is *negative* for all $dX \neq 0$ satisfying $(dX)'X_\circ + X_\circ' dX = 0$, a maximum has been found. The choice $L = (P'P)^{\frac{1}{2}} > 0$ guarantees this.

3 The Green procedure

As stated we use the MSE matrix $M = E\left(\hat{f} - f\right)\left(\hat{f} - f\right)' = \left(a + B'\mu_y\right)\left(a + B'\mu_y\right)' + B'\Omega B + \Phi - B'A\Phi - \Phi A'B$. Obviously $a + B'\mu_y = 0$, as we have to minimize $\mathrm{tr}M$. As a consequence $E\hat{f}\hat{f}' = B'\Omega B$. Imposition of the constraint $E\hat{f}\hat{f}' = Eff'$ yields then $M = 2\Phi - B'A\Phi - \Phi A'B$. Green (1969) defines the problem:

$$\min_B \ \mathrm{tr}\,(2\Phi - B'A\Phi - \Phi A'B) \quad \text{subject to } B'\Omega B = \Phi.$$

We introduce $C' = \Phi^{-\frac{1}{2}}B'\Omega^{\frac{1}{2}}$. Clearly $C'C = I_m$. This yields the equivalent problem

$$\max_C \ \mathrm{tr}\,\Phi^{\frac{3}{2}}A'\Omega^{+\frac{1}{2}}C \quad \text{subject to } C'C = I_m$$

We used: $A'\Omega^{+\frac{1}{2}}C\,\Phi^{\frac{1}{2}} = R'\Omega^{\frac{1}{2}}\Omega^{+\frac{1}{2}}\Omega^{\frac{1}{2}}B = R'\Omega^{\frac{1}{2}}B = A'B$, with $A = \Omega^{\frac{1}{2}}R$ due to (1.1).

Application of the Kristof-type theorem gives the solution

$$C_G = \Omega^{+\frac{1}{2}}A\Phi^{\frac{3}{2}}\left(\Phi^{\frac{3}{2}}A'\Omega^{+}A\Phi^{\frac{3}{2}}\right)^{-\frac{1}{2}},$$

from which follows the solution

$$B_G = \Omega^{+}A\Phi^{\frac{3}{2}}\left(\Phi^{\frac{3}{2}}A'\Omega^{+}A\Phi^{\frac{3}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}} + \left(I_p - \Omega^{+\frac{1}{2}}\Omega^{\frac{1}{2}}\right)Q, \quad Q \text{ arbitrary.}$$

The arbitrary component disappears in the predictor expression $B_G'(y - \mu_y)$, because $\left(I_p - \Omega^{\frac{1}{2}}\Omega^{+\frac{1}{2}}\right)\left(y - \mu_y\right) = 0$ with probability one (w. p. 1).

Hence we get as predictor

$$\hat{f}_G = \Phi^{\frac{1}{2}}\left(\Phi^{\frac{3}{2}}A'\Omega^{+}A\Phi^{\frac{3}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{3}{2}}A'\Omega^{+}\left(y - \mu_y\right).$$

The reader can verify that $A'\Omega^{+}A > 0$.

An alternative expression is

$$C_G = F_2G_2',$$

where we have used the singular-value decomposition

$$\Omega^{+\frac{1}{2}}A\Phi^{\frac{3}{2}} = F_2\Gamma_2^{\frac{1}{2}}G_2',$$

with $F_2'F_2 = G_2'G_2 = G_2G_2' = I_m$. Use was made of the fact that $\Omega^{+\frac{1}{2}}A$ has full column rank (m).

For nonsingular Ω the solution becomes that given by ten Berge *et al.* (1999) in their presentation, namely between (6) and (7).

4 The McDonald procedure

This approach is based on the weighted-least-squares function

$$\operatorname{tr}\Psi^+ E\left(y-\mu_y-A\hat{f}\right)\left(y-\mu_y-A\hat{f}\right)'.$$

Clearly

$$E\left(y-\mu_y-A\hat{f}\right)\left(y-\mu_y-A\hat{f}\right)' = \left(I_p-AB'\right)\Omega\left(I_p-BA'\right)+ \\ +A\left(a+B'\mu_y\right)\left(a+B'\mu_y\right)'A'.$$

Again we find that $a+B'\mu_y=0$, now having to minimize

$$\operatorname{tr}\Psi^+ E\left(y-\mu_y-A\hat{f}\right)\left(y-\mu_y-A\hat{f}\right)'.$$

Notice that $A'\Psi^+A>0$.

Imposition of the constraint $E\hat{f}\hat{f}' = Eff'$ leads to the problem of minimizing

$$\operatorname{tr}\Psi^+\left(I_p-AB'\right)\Omega\left(I_p-BA'\right) \quad \text{subject to } B'\Omega B=\Phi.$$

Using $C'=\Phi^{-\frac{1}{2}}B'\Omega^{\frac{1}{2}}$ we define the problem:

$$\max_C \operatorname{tr}\Phi^{\frac{1}{2}}A'\Psi^+\Omega^{\frac{1}{2}}C \quad \text{subject to } C'C=I_m.$$

Application of the Kristof-type theorem yields the solution

$$C_M=\Omega^{\frac{1}{2}}\Psi^+A\Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Psi^+\Omega\Psi^+A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}},$$

from which follows the solution

$$B_M=\Omega^{+\frac{1}{2}}\Omega^{\frac{1}{2}}\Psi^+A\Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Psi^+\Omega\Psi^+A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}}+\left(I_p-\Omega^{+\frac{1}{2}}\Omega^{\frac{1}{2}}\right)Q, \quad Q \text{ arbitrary.}$$

Finally the predictor turns out to be

$$\hat{f}_M=\Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Psi^+\Omega\Psi^+A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}}A'\Psi^+\left(y-\mu_y\right).$$

Again we used

$$\left(I_p-\Omega^{\frac{1}{2}}\Omega^{+\frac{1}{2}}\right)\left(y-\mu_y\right)=0 \quad \text{w.p.1.}$$

The reader can verify that $A'\Psi^+\Omega\Psi^+A>0$, using

$$A'\Psi^+\Omega\Psi^+A=A'\Psi^+\left(A\Phi A'+\Psi\right)\Psi^+A=A'\Psi^+A\Phi A'\Psi^+A+A'\Psi^+A.$$

An alternative expression is

$$C_M = F_3 G_3',$$

where

$$\Omega^{\frac{1}{2}}\Psi^{+}A\Phi^{\frac{1}{2}} = F_3\,\Gamma_3^{\frac{1}{2}} G_3',$$

with $F_3'F_3 = G_3'G_3 = G_3G_3' = I_m$.

For nonsingular Ω the solution becomes that given by ten Berge *et al.* (1999) in their presentation, namely between (4) and (5).

5 The Krijnen *et al.* procedure

Like Green's this approach uses the MSE matrix M of $\hat{f}$. Instead of $\operatorname{tr}(2\Phi - B'A\Phi - \Phi A'B)$, Krijnen *et al.* use $|2\Phi - B'A\Phi - \Phi A'B|$ which has to be minimized. The first thing to do is to prove that $2\Phi - B'A\Phi - \Phi A'B > 0$.

We have

$$\begin{aligned} 2\Phi - B'A\Phi - \Phi A'B &= \Phi^{\frac{1}{2}}\left(2I_m - \Phi^{-\frac{1}{2}}B'A\Phi^{\frac{1}{2}} - \Phi^{\frac{1}{2}}A'B\Phi^{-\frac{1}{2}}\right)\Phi \\ &= \Phi^{\frac{1}{2}}\left(2I_m - \Phi^{-\frac{1}{2}}B'\Omega^{\frac{1}{2}}\Omega^{+\frac{1}{2}}A\Phi^{\frac{1}{2}} - \Phi^{\frac{1}{2}}A'\Omega^{+\frac{1}{2}}\Omega^{\frac{1}{2}}B\Phi^{-\frac{1}{2}}\right)\Phi^{\frac{1}{2}} \\ &= \Phi^{\frac{1}{2}}\left(2I_m - C'V - V'C\right)\Phi^{\frac{1}{2}} \\ &= \Phi^{\frac{1}{2}}\left[(C-V)'(C-V) + (I_m - V'V)\right]\Phi^{\frac{1}{2}} \end{aligned}$$

where

$$V = \Omega^{+\frac{1}{2}}A\Phi^{\frac{1}{2}} \quad \text{and hence} \quad V'V = \Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}}.$$

We shall show that all eigenvalues of $V'V$ are positive and less than unity. Pre-(post-) multiply the moment structure $\Omega = A\Phi A' + \Psi$ by $\Phi^{\frac{1}{2}}A'\Omega^{+}\left(\Omega^{+}A\Phi^{\frac{1}{2}}\right)$. This leads to $\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}} = \left(\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}}\right)^2 + \Phi^{\frac{1}{2}}A'\Omega^{+}\Psi\Omega^{+}A\Phi^{\frac{1}{2}}$, hence

$$\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}} > \left(\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}}\right)^2,$$

as $\Phi^{\frac{1}{2}}A'\Omega^{+}\Psi\Omega^{+}A\Phi^{\frac{1}{2}} > 0$, and $\lambda_i > \lambda_i^2$ where λ_i is any eigenvalue of $\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}}$. This proves the property. Hence $I_m - V'V > 0$. As $(C-V)'(C-V) \geqslant 0$ we have shown that $2\Phi - B'A\Phi - \Phi A'B > 0$.

Hence $|2\Phi - B'A\Phi - \Phi A'B| > 0$. Consider then the positive definite matrix $2I_m - C'V - V'C$. We use (7.18) in Zhang (1999) which yields

$$C'V + V'C \leqslant 2U'\left(V'CC'V\right)^{\frac{1}{2}}U$$

where U is an orthogonal matrix.

As $C'C = I_m$ we have $CC' \leqslant I_p$. This in its turn leads to $V'CC'V \leqslant V'V$. The latter inequality gives $(V'CC'V)^{\frac{1}{2}} \leqslant (V'V)^{\frac{1}{2}}$. See Theorem 2.5.5 in Wang & Chow (1994).

Finally, we have

$$C'V + V'C \leqslant 2U'(V'V)^{\frac{1}{2}}U$$

or equivalently

$$2I_m - C'V - V'C \geqslant 2\left[I_m - U'(V'V)^{\frac{1}{2}}U\right].$$

From this we derive

$$\left|2I_m - C'V - V'C\right| \geqslant \left|2\left[I_m - U'(V'V)^{\frac{1}{2}}U\right]\right| = \left|2\left[I_m - (V'V)^{\frac{1}{2}}\right]\right|.$$

It is easy to see that $C_K = V(V'V)^{-\frac{1}{2}}$ leads to the equality

$$\left|2I_m - C_K'V - V'C_K\right| = \left|2\left[I_m - (V'V)^{\frac{1}{2}}\right]\right|.$$

Hence C_K solves the problem. It is not clear whether the solution is unique.

In fact, C_K also solves the related problem

$$\max_C \operatorname{tr} V'C \quad \text{subject to } C'C = I_m.$$

The (unique) solution is C_K by the Kristof-type theorem.

Application of Zhang's (7.18) yields

$$2\operatorname{tr} V'C = \operatorname{tr}(C'V + V'C) \leqslant 2\operatorname{tr} U'(V'V)^{\frac{1}{2}}U = 2\operatorname{tr}(V'V)^{\frac{1}{2}},$$

which again has solution C_K. We then get the solution

$$B_K = \Omega^+ A\Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Omega^+ A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}} + \left(I_p - \Omega^{+\frac{1}{2}}\Omega^{\frac{1}{2}}\right)Q, \quad Q \text{ arbitrary.}$$

From this follows the *unique* predictor

$$\hat{f}_K = \Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Omega^+ A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}}A'\Omega^+\left(y - \mu_y\right).$$

For nonsingular Ω the solution C_K coincides with that given by ten Berge *et al.* (1999), namely in (9).

6 Equality of $\hat{f}_G$ and $\hat{f}_K$ when Φ and $A'\Omega^+A$ commute

ten Berge *et al.* (1999) showed that $C_G = C_K$ under their assumptions when Φ and $A'\Omega^{-1}A$ commute. We shall prove that $\hat{f}_G = \hat{f}_K$ under our milder conditions.

When Φ and $A'\Omega^+A$ commute we have $\Phi = SMS'$ and $A'\Omega^+A = SNS'$, where M and N are positive definite diagonal matrices and S is orthogonal. Hence

$$\begin{aligned}\Phi^{\frac{3}{2}}\left(\Phi^{\frac{3}{2}}A'\Omega^+A\Phi^{\frac{3}{2}}\right)^{-\frac{1}{2}} &= SM^{\frac{3}{2}}S'\left(SM^{\frac{3}{2}}S'SNS'SM^{\frac{3}{2}}S'\right)^{-\frac{1}{2}}\\ &= SM^{\frac{3}{2}}S'\left(SM^3NS'\right)^{-\frac{1}{2}} = SM^{\frac{3}{2}}S'S\left(M^3N\right)^{-\frac{1}{2}}S'\\ &= SN^{-\frac{1}{2}}S' = (A'\Omega^+A)^{-\frac{1}{2}}.\end{aligned}$$

Further $\Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}} = (A'\Omega^{+}A)^{-\frac{1}{2}}$.

This yields $\hat{f}_G = \hat{f}_K = \Phi^{\frac{1}{2}}(A'\Omega^{+}A)^{-\frac{1}{2}}A'\Omega^{+}\left(y-\mu_y\right)$.

7 Equality of $\hat{f}_M$ and $\hat{f}_K$ when Ψ and $A\Phi A'$ commute

ten Berge *et al.* (1999) showed that $C_M = C_K$ under their assumptions when Ψ is nonsingular. Essential is the expression

$$\Omega^{-1} = \Psi^{-1} - \Psi^{-1}A\Phi^{\frac{1}{2}}\left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{-1}A\Phi^{\frac{1}{2}}\right)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{-1}.$$

Under our assumptions $I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}$ is nonsingular because $A'\Psi^{+}A > 0$ which follows from $T'\Psi T > 0$ and (1.1). When we additionally assume that Ψ and $A\Phi A'$ commute we can establish the equality

$$\Omega^{+} = \Psi^{+} - \Psi^{+}A\Phi^{\frac{1}{2}}\left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+}.$$

Proof. When Ψ and $A\Phi A'$ commute we have $\Psi = SMS'$ and $A\Phi A' = SNS'$ where M and N are positive definite diagonal matrices and $S'S = I_m$. Further $A\Phi^{\frac{1}{2}} = SN^{\frac{1}{2}}T'$, with orthogonal T, a singular-value decomposition. Hence

$$\begin{aligned}
&\Psi^{+} - \Psi^{+}A\Phi^{\frac{1}{2}}\left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+}\\
&\quad = SM^{-1}S' - SM^{-1}S'SN^{\frac{1}{2}}T'\left(I_m + TN^{\frac{1}{2}}S'SM^{-1}S'SN^{\frac{1}{2}}T'\right)^{-1}TN^{\frac{1}{2}}S'SM^{-1}S'\\
&\quad = SM^{-1}S' - SM^{-1}N^{\frac{1}{2}}T'\left(I_m + TM^{-1}NT'\right)^{-1}TM^{-1}N^{\frac{1}{2}}S'\\
&\quad = SM^{-1}S' - SM^{-1}N^{\frac{1}{2}}T'T\left(I_m + M^{-1}N\right)^{-1}T'TM^{-1}N^{\frac{1}{2}}S'\\
&\quad = SM^{-1}S' - SM^{-1}N^{\frac{1}{2}}\left(I_m + M^{-1}N\right)^{-1}M^{-1}N^{\frac{1}{2}}S'.
\end{aligned}$$

Further $\Omega = A\Phi A' + \Psi = S(M+N)S'$, and $\Omega^{+} = S(M+N)^{-1}S'$. It is easy to see that

$$(M+N)^{-1} = M^{-1} - M^{-1}N^{\frac{1}{2}}\left(I_m + M^{-1}N\right)^{-1}M^{-1}N^{\frac{1}{2}}.$$

This yields the result. □

Recall that

$$\hat{f}_M = \Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Psi^{+}\Omega\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}}A'\Psi^{+}(y-\mu_y)$$

and

$$\hat{f}_K = \Phi^{\frac{1}{2}}\left(\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}}\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}}A'\Omega^{+}(y-\mu_y).$$

Consider

$$\begin{aligned}
\Phi^{\frac{1}{2}}A'\Omega^{+}A\Phi^{\frac{1}{2}} &= \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}} - \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-1} \\
&\times \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}} \\
&= \left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}} \\
&= (I_m + E)^{-1}E, \\
\Phi^{\frac{1}{2}}A'\Psi^{+}\Omega\Psi^{+}A\Phi^{\frac{1}{2}} &= \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi A'\Psi^{+}A\Phi^{\frac{1}{2}} + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}} \\
&= \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}} + \left(\Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^2 \\
&= E + E^2, \\
\Phi^{\frac{1}{2}}A'\Omega^{+} &= \Phi^{\frac{1}{2}}A'\Psi^{+} - \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+} \\
&= \left(I_m + \Phi^{\frac{1}{2}}A'\Psi^{+}A\Phi^{\frac{1}{2}}\right)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+} \\
&= (I_m + E)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+},
\end{aligned}$$

$$\begin{aligned}
\hat{f}_K &= \Phi^{\frac{1}{2}}\left[(I_m + E)^{-1}E\right]^{-\frac{1}{2}}(I_m + E)^{-1}\Phi^{\frac{1}{2}}A'\Psi^{+}\left(y - \mu_y\right), \\
\hat{f}_M &= \Phi^{\frac{1}{2}}\left(E + E^2\right)^{-\frac{1}{2}}\Phi^{\frac{1}{2}}A'\Psi^{+}\left(y - \mu_y\right).
\end{aligned}$$

Clearly

$$\left[(I_m + E)^{-1}E\right]^{-\frac{1}{2}}(I_m + E)^{-1} = \left(E + E^2\right)^{-\frac{1}{2}} \quad \text{as } E > 0.$$

This establishes the equality of $\hat{f}_K$ and $\hat{f}_M$.

8 Comments

1. ten Berge *et al.* (1999) claim that the McDonald method is undefined when Ψ is singular. This is unjustified. What matters is the nonsingularity of $T'\Psi T$. We make that assumption. It implies that $A'\Psi^{+}A > 0$ which we use several times.
2. Application of Zhang's result shows immediately that C_G and C_M yield the maximum. The Kristof-type theorem shows the *unicity* of the solutions.

Acknowledgement

The author is grateful to Götz Trenkler for drawing his attention to Zhang's result (7.18) which was fruitfully applied in Section 5. The reasoning why $B'\mu_y + a$ should be equal

to zero in all three procedures is due to Albert Satorra. This yields $\hat{f} = B'\left(y - \mu_y\right)$, an unbiased predictor.

9 References

Anderson, T. W. & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. Vol. V, 111-50. University of California Press, Berkeley and Los Angeles.

ten Berge, J. M. F. (1993). *Least Squares Optimization in Multivariate Analysis*. *DSWO Press*, Leiden, Netherlands.

ten Berge, J. M. F., Krijnen, W. P., Wansbeek, T. & Shapiro, A. (1999). Some new results on correlation-preserving factor scores prediction methods. *Linear Algebra Appl.*, 289, 311-8.

Green, B. F. (1969). Best linear composites with a specified structure. *Psychometrika*, 34, 301-18.

Krijnen, W. P., Wansbeek, T. J. & ten Berge, J. M. F. (1996). Best linear predictors for factor scores. *Communications in Statistics: Theory and Methods*, 25, 3013-25.

Kristof, W. (1970). A theorem on the trace of certain matrix products and some applications. *Journal of Mathematical Psychology*, 7, 515-30.

Löwner, K. (1934). Über monotone Matrixfunktionen. *Mathematische Zeitschrift*, 38, 177-216.

Magnus, J. R. & Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics. Revised Edition*. Wiley, Chichester, England.

McDonald, R. P. (1981). Constrained least squares estimators of oblique common factors. *Psychometrika*, 46, 337-41.

Wang, S. G. & Chow, S. C. (1994). *Advanced Linear Models: Theory and Applications*. Marcel Dekker, New York.

Zhang, F. (1999). *Matrix Theory. Basic Results and Techniques*. Springer, New York.

Resum

Es dóna una generalització dels resultats presentats per ten Berge, Krijnen, Wansbeek and Shapiro . Aquests autors examinen mètodes i resultats basats en Anderson i Rubin. Mc Donald, Green i Krijnen, Wansbeek i ten Berge. Considerarem el mateix plantejament però sota condicions de rang més dèbils. Així suposarem que alguns moments, com les matrius de covariàncies Ω del vector de mesures observades dels factors comuns i ψ dels factors únics, siguin singulars. Imposem la condició $T'\psi T > 0$, essent $T\Lambda T'$ la descomposició de Schur de Ω. Com és usual, suposem que tenen rang màxim per columnes les matrius de covariàncies Φ dels factors comuns i la matriu A del model factorial.

MSC: 62H25, 15A24

Paraules clau: anàlisi factorial, mesures de factors, preservació de la covariància, teorema tipus Kristof