## A Supplementary Note on the Paper:

"Oblique Factors and Components with Independent Clusters"

## Haruhiko Ogasawara

This note is to supplement Ogasawara (2003). The equation numbers without the prefix 'B' in this article are those of Ogasawara (2003).

## 1. Equivalence of the unweighted least squares estimator and the Wishart maximum likelihood estimator

For the derivation of the equivalence of the unweighted least squares (ULS) estimator and the maximum Wishart likelihood (WML) estimator assuming multivariate normality for the direct-product model of (28a), we give the following Lemma A1.

**Lemma A1.** Let  $\Sigma = \Sigma(\theta)$  be a  $q \times q$  population covariance matrix for observed variables with  $\theta = (\theta_1, ..., \theta_t)'$  being a  $t \times 1$  parameter vector,  $F_{\text{ULS}} = (1/2) \text{tr}\{(\Sigma - \mathbf{S})^2\}$ , and  $F_{\text{WML}} = \text{tr}(\Sigma^{-1}\mathbf{S}) + \ln |\Sigma|$ , where  $\mathbf{S}$  is a  $q \times q$  sample covariance matrix. Then, a necessary condition for the equivalence of a ULS estimator  $(\hat{\boldsymbol{\theta}}_{\text{ULS}})$  and a WML estimator  $(\hat{\boldsymbol{\theta}}_{\text{WML}})$  is that we have the same  $\hat{\boldsymbol{\theta}}$  which satisfies

$$\frac{\partial F_{\text{ULS}}}{\partial \theta_i}\Big|_{\theta=\hat{\theta}} = \text{tr}\left\{ (\mathbf{\Sigma} - \mathbf{S}) \frac{\partial \mathbf{\Sigma}}{\partial \theta_i} \right\} = 0\Big|_{\theta=\hat{\theta}}, (i = 1, ..., t)$$
(B1)

and

$$\frac{\partial F_{\text{WML}}}{\partial \theta_{i}}\Big|_{\theta=\hat{\theta}} = -\text{tr}\left\{ (\mathbf{\Sigma} - \mathbf{S}) \frac{\partial \mathbf{\Sigma}^{-1}}{\partial \theta_{i}} \right\} = 0\Big|_{\theta=\hat{\theta}}, (i = 1, ..., t).$$
(B2)

Proof. The necessary condition of (B1) for  $\hat{\boldsymbol{\theta}}_{ULS}$  is well-known. The corresponding condition for  $\hat{\boldsymbol{\theta}}_{WML}$  is also well-known and is usually written as

$$\operatorname{tr}\left\{\mathbf{\Sigma}^{-1}(\mathbf{\Sigma}-\mathbf{S})\mathbf{\Sigma}^{-1}\partial\mathbf{\Sigma}/\partial\boldsymbol{\theta}_{i}\right\}=0|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\text{unif}}}, (i=1,...,t)$$

(see e.g., Jöreskog, 1978) which is equivalent to (B2) from  $\partial \Sigma^{-1}/\partial \theta_i = -\Sigma^{-1}(\partial \Sigma^{-1}/\partial \theta_i)\Sigma^{-1}$ . The necessary condition of the equivalence of the two estimators follows immediately. Q.E.D.

Now, we have the main result.

**Proposition A1.** For the direct-product covariance structure model of (28a) with associated assumptions, the WML and ULS estimators are the same.

Proof. For ease of derivation, we reparameterize the model into the following linear model with a new set of parameters  $\{\alpha, \beta, \psi\}$ :

$$\Sigma = \mathbf{E}_{q} \lambda^{2} \Phi \mathbf{E}_{q}' + \psi \mathbf{I}_{q}$$

$$= \mathbf{E}_{q} \Gamma \mathbf{E}_{q}' + \psi \mathbf{I}_{q}$$

$$= \alpha \mathbf{1}_{q} \mathbf{1}_{q}' + \beta (\mathbf{I}_{r} \otimes \mathbf{1}_{p} \mathbf{1}_{p}') + \psi \mathbf{I}_{q},$$
(B3)

where

$$\mathbf{E}_{q} = \mathbf{I}_{r} \otimes \mathbf{1}_{p, \text{ and }} \Gamma = \lambda^{2} \Phi = \alpha \mathbf{1}_{r} \mathbf{1}_{r} + \beta \mathbf{I}_{r}$$
(B4)

with  $\alpha = \lambda^2 \phi$  and  $\beta = \lambda^2 (1 - \phi)$ . Using the reparameterization of (B3) and Lemma A1, the ULS estimators of  $\alpha, \beta$  and  $\psi$  are obtained as the solution of the following equations:

$$\frac{\partial F_{\text{ULS}}}{\partial (\alpha, \beta, \psi)'} = \begin{bmatrix} \text{tr}\{(\mathbf{\Sigma} - \mathbf{S})\mathbf{1}_{q}\mathbf{1}_{q}'\} \\ \text{tr}\{(\mathbf{\Sigma} - \mathbf{S})(\mathbf{I}_{r} \otimes \mathbf{1}_{p}\mathbf{1}_{p}')\} \\ \text{tr}(\mathbf{\Sigma} - \mathbf{S}) \end{bmatrix} = \mathbf{0}.$$
 (B5)

For the WML estimators of  $\alpha, \beta$  and  $\psi$ , we first obtain  $\Sigma^{-1}$ 

by using the formula for the inverse of the sum of two matrices:

$$\Sigma^{-1} = \frac{1}{\psi} \mathbf{I}_{q} - \frac{1}{\psi^{2}} \mathbf{E}_{q} (\mathbf{E}_{q}' \frac{1}{\psi} \mathbf{E}_{q} + \Gamma^{-1})^{-1} \mathbf{E}_{q}'$$

$$= \frac{1}{\psi} \mathbf{I}_{q} - \frac{1}{\psi} \mathbf{E}_{q} (p \mathbf{I}_{r} + \psi \Gamma^{-1})^{-1} \mathbf{E}_{q}'.$$
(B6)

Since the matrix  $\Gamma^{-1}$  in (B5) is obtained as

$$\Gamma^{-1} = \frac{1}{\beta} (\mathbf{I}_r - \frac{\alpha}{\beta + r\alpha} \mathbf{1}_r \mathbf{1}_r'), \tag{B7}$$

we have

$$(p\mathbf{I}_{r} + \psi \mathbf{\Gamma}^{-1})^{-1} = \left\{ \left( p + \frac{\psi}{\beta} \right) \mathbf{I}_{r} - \frac{\alpha \psi}{\beta(\beta + r\alpha)} \mathbf{1}_{r} \mathbf{1}_{r}' \right\}^{-1}$$

$$= \frac{1}{p + \frac{\psi}{\beta}} \left( \mathbf{I}_{r} + \frac{\frac{\alpha \psi}{\beta(\beta + r\alpha)}}{p + \frac{\psi}{\beta} - \frac{r\alpha \psi}{\beta(\beta + r\alpha)}} \mathbf{1}_{r} \mathbf{1}_{r}' \right)^{(B8)}$$

and consequently

$$\Sigma^{-1} = g_1 \mathbf{1}_q \mathbf{1}_q' + g_2 (\mathbf{I}_r \otimes \mathbf{1}_p \mathbf{1}_p') + g_3 \mathbf{I}_q$$
 (B9)

with

$$g_1 = \frac{-\alpha}{(p\beta + \psi)(p\beta + \psi + pr\alpha)}, \quad g_2 = \frac{-\beta}{\psi(p\beta + \psi)},$$

$$g_3 = \frac{1}{\psi}, \quad (B10)$$

Using Lemma A1, (B9) and (B10), the WML estimator of the parameters are obtained as the solution of the following equations:

$$\frac{\partial F_{\text{WML}}}{\partial (\alpha, \beta, \psi)'} = -\text{tr}\left\{ (\mathbf{\Sigma} - \mathbf{S}) \mathbf{1}_{q} \mathbf{1}_{q}' \right\} \frac{\partial g_{1}}{\partial (\alpha, \beta, \psi)'} 
- \text{tr}\left\{ (\mathbf{\Sigma} - \mathbf{S}) (\mathbf{I}_{r} \otimes \mathbf{1}_{p} \mathbf{1}_{p}') \right\} \frac{\partial g_{2}}{\partial (\alpha, \beta, \psi)'} 
- \text{tr}\left\{ (\mathbf{\Sigma} - \mathbf{S}) \right\} \frac{\partial g_{3}}{\partial (\alpha, \beta, \psi)'}$$

$$= \mathbf{0}. \tag{B11}$$

From (B5) and (B11), we find that the ULS estimators satisfy (B11) and consequently that the ULS estimators are also the WML estimators. Since the reparameterized model has identification, the ULS and WML estimators are the same. O.E.D.

The derivation when r=1, which reduces to the case of the one-factor model for parallel tests, was given by Ogasawara (1990) by explicitly deriving the WML and ULS estimators.

Let a covariance structure model be identified and linear in parameters as  $\Sigma = \sum_{i=1}^{t} \theta_{i} \mathbf{K}_{i}$ , where  $\mathbf{K}_{i}$ , (i=1,...,t) are given or design matrices. Then if  $\Sigma^{-1}$  is described as

$$\mathbf{\Sigma}^{-1} = \sum_{i=1}^{t} g_i(\mathbf{\theta}) \mathbf{K}_i, \tag{B12}$$

where  $g_i(\theta)$ , (i=1,...,t) are scalar functions of  $\theta$ , we see that the WML and ULS estimators of the parameters in the covariance structure model are the same. The model of (B12) is a linear model of  $\Sigma^{-1}$  with respect to  $g_i(\theta)$ 's (for estimation and testing in this model, see Anderson, 1969). Other than the model of (28a) or its special cases (e.g., the parallel test model when r=1), we have the model with the form of (B12) when e.g.,  $\Sigma = \text{diag}(\sigma_{11},...,\sigma_{qq})$ , which represents independent observed variables with unconstrained

variances, and is used as a baseline model to assess the goodness-of-fit of covariance structure models (see e.g., Bentler & Bonett, 1980). For the baseline model, it is easy to have  $\Sigma^{-1}$  and the WML and ULS estimators become  $\hat{\sigma}_{ii} = S_{ii}$ , (i = 1, ..., t), where  $S_{ii}$  is the sample variance of the *i*-th observed variable (see e.g., Weng & Cheng, 1997).

The generalized least squares (GLS) estimators are also used in structural equation modeling. It is known that the GLS estimators are obtained as the solutions of the following equations:

$$\operatorname{tr}\left\{\mathbf{S}^{-1}(\mathbf{\Sigma}-\mathbf{S})\mathbf{S}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \boldsymbol{\theta}_{i}}\right\}=0, (i=1,...,t)$$
 (B13)

which is equivalent to

$$\operatorname{tr}\left\{ (\mathbf{\Sigma} - \mathbf{S}) \frac{\partial \mathbf{S}^{-1} \mathbf{\Sigma} \mathbf{S}^{-1}}{\partial \theta_{i}} \right\} = 0, (i = 1, ..., t)$$
 (B14)

For a linear model with the same WML and ULS estimators, (B13) or (B14) are not necessarily satisfied except for e.g., the special case when S is proportional to  $I_q$ . That is, the GLS estimators for the direct-product model of (28a) are generally different from those of the WML and ULS estimators though they are asymptotically equivalent. However, it is to be noted that the GLS estimators for a linear model can be explicitly obtained (the author is indebted to a reviewer for this point).

## 2. The derivations of the asymptotic variances and covariances

$$\mathbf{2.1} \quad \operatorname{avar}(s_{\text{off}}) = \operatorname{avar}(\sum_{i \neq j} \mathbf{1}_p ' \mathbf{S}_{ij} \mathbf{1}_p)$$

We derive 
$$\operatorname{avar}(\sum_{i\neq j} \mathbf{1}_p \mathbf{S}_{ij} \mathbf{1}_p)$$
 as  $\operatorname{4avar}(\sum_{i>j} \mathbf{1}_p \mathbf{S}_{ij} \mathbf{1}_p)$ .

Figure A shows a schematic representation of S with p=3 and r=4.

				-						
		•				1				
5	5	5								
4	4	4								
5	5	5			•					
5	5	5	3	2	3					
4	4	4	2	1	2					
5	5	5	3	2	3					
6	6	6	5	4	5	5	4	5		
6	6	6	5	4	5	5	4	5	•	
6	6	6	5	4	5	5	4	5	•	•

Figure A. Schematic representation for  $\operatorname{avar}(\sum_{i>j}\mathbf{1}_p'\mathbf{S}_{ij}\mathbf{1}_p)$  when p=3 and r=4.

The elements denoted by dots are not included in the off-diagonal blocks  $\mathbf{S}_{ij}$ ,  $(r \ge i > j \ge 1)$ . We have six different asymptotic variances/covariances between the elements of  $\mathbf{S}_{ij}$ ,  $(r \ge i > j \ge 1)$ , which are schematically denoted by "1" through "6" in the figure. We consider an element denoted by "1" with which the asymptotic covariances of the elements in  $\mathbf{S}_{ij}$ ,  $(r \ge i > j \ge 1)$  are to be evaluated. The figure shows an example when the element denoted by "1" is  $(\mathbf{S}_{32})_{22}$ . The following show the six covariance types corresponding to the elements denoted by "1" through "6" in the figure, their values and the numbers of the associated elements for a particular element denoted by "1" in the figure.

Type Value Number of the elements

(1) avar(
$$(\mathbf{S}_{ij})_{ab}$$
), 
$$(r \ge i > j \ge 1; \ a,b = 1,...,p)$$
Number of the elements

(2) 
$$\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{ij})_{ac}),$$
 $\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{ij})_{cb}),$ 
 $(r \ge i > j \ge 1;$ 
 $a,b,c = 1,...,p;$ 
 $b \ne c,a \ne c)$ 
(3)  $\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{ij})_{cd}),$ 
 $(r \ge i > j \ge 1;$ 
 $a,b,c,d = 1,...,p;$ 
 $a \ne c,b \ne d)$ 
(4)  $\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{ij})_{cb}),$ 
 $(r \ge i > j \ge 1; r \ge i > k \ge 1; r \ge 1; r \ge 1$ 

Noting that the number of the elements in  $\mathbf{S}_{ij}$ ,  $(r \ge i > j \ge 1)$  is  $r(r-1)p^2/2$ , we have  $\operatorname{avar}(\sum_{i \ne j} \mathbf{1}_p ' \mathbf{S}_{ij} \mathbf{1}_p)$ 

$$= \frac{4}{N} \times \frac{r(r-1)p^2}{2} \Big[ (\lambda^2 + \psi)^2 + \lambda^4 \phi^2 + 2(p-1)\{(\lambda^2 + \psi)\lambda^2 + \lambda^4 \phi^2\} \\ + (p^2 - 2p + 1)(\lambda^4 + \lambda^4 \phi^2) + 2(r-2)p\{(\lambda^2 + \psi)\lambda^2 \phi + \lambda^4 \phi^2\} \\ + 2(r-2)(p^2 - p)(\lambda^4 \phi + \lambda^4 \phi^2) + \frac{r^2 - 5r + 6}{2} p^2 (2\lambda^4 \phi^2) \Big]$$

$$= \frac{2r(r-1)p^2}{N} \Big[ \{1 + 2(p-1) + (p^2 - 2p + 1)\}\lambda^4 + \{2 + 2(p-1)\}\lambda^2 \psi \\ + \psi^2 + 2(r-2)p\psi \lambda^2 \phi + \{1 + 2(p-1) + (p^2 - 2p + 1) + 2(r-2)p + 2(r-2)(p^2 - p) + (r^2 - 5r + 6)p^2\}\lambda^4 \phi^2 \\ + \{2(r-2)p + 2(r-2)(p^2 - p)\}\lambda^4 \phi \Big]$$

$$= \frac{2r(r-1)p^2}{N} \Big\{ p^2\lambda^4 + 2p\lambda^2 \psi + \psi^2 + 2(r-2)p\psi \lambda^2 \phi \\ + (r^2 - 3r + 3)p^2\lambda^4 \phi^2 + 2(r-2)p^2\lambda^4 \phi \}.$$

2.2 
$$\operatorname{acov}(s_{\text{off}}, s_{\text{d}} - \operatorname{tr} \mathbf{S}) = \operatorname{acov}(\sum_{i \neq j} \mathbf{1}_{p}' \mathbf{S}_{i,j} \mathbf{1}_{p}, \sum_{i=1}^{r} \mathbf{1}_{p}' \mathbf{S}_{i,i} \mathbf{1}_{p} - \operatorname{tr} \mathbf{S})$$

To derive  $\operatorname{acov}(\sum_{i \neq j} \mathbf{1}_{p}' \mathbf{S}_{i,j} \mathbf{1}_{p}, \sum_{i=1}^{r} \mathbf{1}_{p}' \mathbf{S}_{i,i} \mathbf{1}_{p} - \operatorname{tr} \mathbf{S})$ , w

evaluate three covariance types as follows:

Type Value Number of the elements (1) 
$$acov((\mathbf{S}_{ij})_{ac}, (\mathbf{S}_{ii})_{ab})$$
,  $acov((\mathbf{S}_{ji})_{ac}, (\mathbf{S}_{ii})_{ba})$ ,  $acov((\mathbf{S}_{ji})_{ac}, (\mathbf{S}_{ii})_{ba})$ ,  $acov((\mathbf{S}_{ji})_{ca}, (\mathbf{S}_{ii})_{ba})$ ,  $acov((\mathbf{S}_{ji})_{ca}, (\mathbf{S}_{ii})_{ba})$ ,  $acov((\mathbf{S}_{ji})_{ca}, (\mathbf{S}_{ii})_{ba})$ ,  $acov((\mathbf{S}_{ji})_{ca}, (\mathbf{S}_{ii})_{ab})$ ,  $acov((\mathbf{S}_{ij})_{ac}, (\mathbf{S}_{ii})_{ab})$ ,  $acov((\mathbf{S}_{ji})_{ac}, (\mathbf{S}_{ii})_{ab})$ ,  $acov((\mathbf{S}_{kl})_{cd}, (\mathbf{S}_{ii})_{ab})$ ,  $acov((\mathbf{S}_{ii})_{ab}, (\mathbf{$ 

Then, we have

$$acov(\sum_{i\neq j} \mathbf{1}_{p} \mathbf{S}_{ij} \mathbf{1}_{p}, \sum_{i=1}^{r} \mathbf{1}_{p} \mathbf{S}_{ii} \mathbf{1}_{p} - tr \mathbf{S})$$

$$= \frac{1}{N} \Big[ 4r(r-1)p^{2} (p-1)\{(\lambda^{2} + \psi)\lambda^{2}\phi + \lambda^{4}\phi\} + 2r(r-1)p^{2}(p^{2} - 3p + 2)(2\lambda^{4}\phi) + r(r-1)(r-2)p^{3}(p-1)(2\lambda^{4}\phi^{2}) \Big]$$

$$= \frac{2}{N} r(r-1)p^{2}(p-1)\{2p\lambda^{4}\phi + 2\psi\lambda^{2}\phi + p(r-2)\lambda^{4}\phi^{2}\}.$$

**2.3** 
$$acov(s_{off}, s_d) = acov(\sum_{i \neq j} \mathbf{1}_p ' \mathbf{S}_{ij} \mathbf{1}_p, \sum_{i=1}^r \mathbf{1}_p ' \mathbf{S}_{ii} \mathbf{1}_p)$$

We use the result in the previous section as follows:

$$acov(\sum_{i\neq j} \mathbf{1}_{p} \mathbf{S}_{ij} \mathbf{1}_{p}, \sum_{i=1}^{r} \mathbf{1}_{p} \mathbf{S}_{ii} \mathbf{1}_{p})$$

$$= acov(\sum_{i\neq j} \mathbf{1}_{p} \mathbf{S}_{ij} \mathbf{1}_{p}, \sum_{i=1}^{r} \mathbf{1}_{p} \mathbf{S}_{ii} \mathbf{1}_{p} - \operatorname{tr} \mathbf{S})$$

$$+ acov(\sum_{i\neq j} \mathbf{1}_{p} \mathbf{S}_{ij} \mathbf{1}_{p}, \operatorname{tr} \mathbf{S}).$$

To evaluate the second term in the right-hand side of the above equation, we consider three covariance types as follows:

Type Value Number of the elements 
$$(1)\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{ii})_{aa}),$$
  $\operatorname{acov}((\mathbf{S}_{ji})_{ba}, (\mathbf{S}_{ii})_{aa}),$   $(i, j = 1, ..., r; i \neq j;$   $a, b = 1, ..., p)$   $2r(r-1)p^2$ 

(2) 
$$\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{ii})_{cc}),$$
  
 $\operatorname{acov}((\mathbf{S}_{ji})_{ba}, (\mathbf{S}_{ii})_{cc}),$   
 $(i, j = 1, ..., r; i \neq j;$   
 $a, b, c = 1, ..., p; a \neq c)$   
(3)  $\operatorname{acov}((\mathbf{S}_{ij})_{ab}, (\mathbf{S}_{kk})_{cc}),$   
 $(i, j, k = 1, ..., r;$   
 $i \neq j, i \neq k, j \neq k;$   
 $a, b, c = 1, ..., p)$   
 $2\lambda^{4}\phi^{2}$   
 $r(r-1)p^{2}(r-2)p$   
 $r(r-1)p^{2}(r-2)p^{3}$ 

From above and the result in the previous section,

$$acov(\sum_{i\neq j} \mathbf{1}_{p}' \mathbf{S}_{ij} \mathbf{1}_{p}, \sum_{i=1}^{r} \mathbf{1}_{p}' \mathbf{S}_{ii} \mathbf{1}_{p})$$

$$= \frac{2}{N} r(r-1) p^{2} (p-1) \{ 2p\lambda^{4} \phi + 2\psi\lambda^{2} \phi + p(r-2)\lambda^{4} \phi^{2} \}$$

$$+ \frac{1}{N} \Big[ 2r(r-1) p^{2} \{ 2(\lambda^{2} + \psi)\lambda^{2} \phi \}$$

$$+ 2r(r-1) p^{2} (p-1)(2\lambda^{4} \phi) + r(r-1)(r-2) p^{3} (2\lambda^{4} \phi^{2}) \Big]$$

$$= \frac{2}{N} r(r-1) p^{3} \{ 2p\lambda^{4} \phi + 2\psi\lambda^{2} \phi + p(r-2)\lambda^{4} \phi^{2} \}.$$

References (see also the reference list of Ogasawara, 2003)

Anderson, T. W. (1969). Statistical inference for covariance matrices with linear structure. In P. R. Krishnaiah (Ed.), *Multivariate analysis II* (pp.55-66). New York: Academic Press.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

Ogasawara, H. (1990). Latent covariance structure model. Japanese

- Journal of Behaviormetrics, 18, 13-27. (in Japanese with English abstract)
- Ogasawara, H. (2003). Oblique factors and components with independent clusters. *Psychometrika*. (in press).
- Weng, L. J., & Cheng, C. P. (1997). Why might relative fit indices differ between estimators? *Structural Equation Modeling*, 4, 121-128.