

On Boundary Problems in the Estimation of Linear Structural MEM

Myung-Sang Moon

Boundary problems occurring in the maximum likelihood estimation of the linear structural equation-error MEM are briefly reviewed. Proofs of some theorems are reproduced in easy words. A simple example with one predictor variable is provided to illustrate the usage of theorems.

1. Introduction

In this section, the general measurement error model is defined and some basic notation is introduced. Bold-faced letters denote vectors or matrices and all vectors are column ones. The error-free component of a measurement error model is denoted by

$$\phi_i = \beta_0 + g(\pi_i; \beta_1) + q_i, \quad i = 1, 2, 3, \dots, n, \quad (1)$$

where ϕ_i is the i^{th} error-free response variable, $\pi_i' = [\pi_{i1} \ \pi_{i2} \ \dots \ \pi_{ik}]$ is a k -dimensional vector of error-free predictor variables, and $\beta_1' = [\beta_1 \ \beta_2 \ \dots \ \beta_k]$ is a k -dimensional vector of regression parameters. The term q_i denotes the error in the equation, not in the measurement of the response variable, and the q_i 's are assumed to be iid $(0, \sigma_{qq})$ with $\sigma_{qq} > 0$.

Let $z_i = (y_i, x_i')'$ be the i^{th} $(k+1)$ -dimensional vector of observed response and predictor variables and $\xi_i = (\phi_i, \pi_i')'$ be the i^{th} $(k+1)$ -dimensional vector of error-free response and predictor variables. Then, $z_i = \xi_i + w_i$ where $w_i = (v_i, u_i')'$ is the i^{th} $(k+1)$ -dimensional

Department of Statistics, Yonsei University, Wonju-City, Kangwon-Do, 222-701, Korea.

vector of measurement errors. Let

$$\begin{aligned} \mathbf{m}_{zz} &= \begin{pmatrix} m_{yy} & \mathbf{m}_{xy}^t \\ \mathbf{m}_{xy} & \mathbf{m}_{xx} \end{pmatrix} \\ &= (n-1)^{-1} \sum_{i=1}^n (z_i - \bar{z})(z_i - \bar{z})', \text{ where } \bar{z} = n^{-1} \sum_{i=1}^n z_i. \end{aligned}$$

Assume that the measurement errors w_i are iid $(\mathbf{0}, \Sigma_{ww})$, where Σ_{ww} is positive semidefinite (p.s.d). In this work, w_i 's are assumed to follow a multivariate normal distribution. The model defined in (1) is called an equation-error model. When the equation error term q_i is not included, the measurement error model is called a no-equation-error model. If the error-free values π_i are taken to be fixed values, the model is said to be a functional model and if the error-free values π_i are random vectors, the model is said to be a structural model. And, if $g(\pi_i; \beta_1)$ is nonlinear in π_i when β_1 is fixed, or if $g(\pi_i; \beta_1)$ is nonlinear in β_1 when π_i is fixed, then model (1) is defined as a nonlinear model.

Linear structural equation-error model is assumed in this work. With this model, ML estimation procedure is discussed in the next section. Some problems associated with boundary of the parameter space in ML estimation is also discussed with Theorems. Example is provided in the last part of section, using linear MEM with single predictor variable.

I. MLE in Linear Structural MEM with Measurement Error Variances Known

True and observed linear structural equation-error model with k predictors is given by the equation (2).

$$\begin{aligned} \psi_i &= \beta_0 + \pi_i \beta_1 + q_i, \\ y_i &= \beta_0 + \pi_i \beta_1 + v_i, \quad x_i = \pi_i + u_i, \end{aligned} \quad (2)$$

where $v_i = a_i + q_i$, (Note that a_i is the error contained in the measurement of y_i) and

$$\begin{bmatrix} \pi_i \\ v_i \\ u_i \end{bmatrix} \sim MVN \left(\begin{bmatrix} \mu_\pi \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\pi\pi} & \mathbf{0}' & \mathbf{0} \\ \mathbf{0} & \sigma_{vv} & \Sigma_{iv}^t \\ \mathbf{0} & \Sigma_{uv} & \Sigma_{uu} \end{bmatrix} \right).$$

Throughout in this work, Σ_{uu} and Σ_{uv} are assumed to be known. But σ_{vv} is assumed to

be unknown. It is a reasonable assumption since σ_{qq} is usually unknown, even if σ_{aa} is known.

From the log likelihood function adjusted for degrees of freedom and the functional invariance property of the method of maximum likelihood, the ML estimators adjusted for degrees of freedom are obtained as given in equation (3).

$$\begin{aligned} \hat{\beta}_1 &= (\mathbf{m}_{xx} - \Sigma_{uu})^{-1} (\mathbf{m}_{xy} - \Sigma_{uv}) & \hat{\beta}_0 &= \bar{y} - \bar{x}' \hat{\beta}_1, \\ \hat{\sigma}_{vv} &= m_{yy} - (\mathbf{m}_{xy} - \Sigma_{uv})' (\mathbf{m}_{xx} - \Sigma_{uu})^{-1} (\mathbf{m}_{xy} - \Sigma_{uv}), \\ &\text{and } \hat{\Sigma}_{\pi\pi} &= \mathbf{m}_{xx} - \Sigma_{uu} \end{aligned} \tag{3}$$

But, since $\Sigma_{\pi\pi}$ is positive definite (p.d.) by assumption and $\sigma_{vv} > 0$, their estimators, to be reasonable ones, should satisfy some conditions. They are stated in Theorem 2.1 and Theorem 2.2 after one useful Result, which is necessary in the following steps, is introduced.

Result 2.1 Let **A** be a symmetric $p \times p$ matrix and let **B** be a symmetric p.d. $p \times p$ matrix. Then the roots, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, of

$$| \mathbf{A} - \lambda \mathbf{B} | = 0,$$

are called the roots of **A** in the metric **B**. There exists a matrix **T** such that

$$\mathbf{T}' \mathbf{B} \mathbf{T} = \mathbf{I} \text{ and } \mathbf{T}' \mathbf{A} \mathbf{T} = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p).$$

Columns of **T** are called the characteristic vectors of **A** in the metric **B**. ■

Theorem 2.1 Suppose that model (2) holds. Then, estimators given in (3) are MLE provided $\hat{\Sigma}_{\pi\pi}$ is positive definite and $\hat{\sigma}_{vv} \geq \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}$, where Σ_{uu}^+ is the Moore-Penrose generalized inverse of Σ_{uu} . If either of these conditions is not met, the estimators fall on the boundary of the parameter space. Note that the formulation given in Theorem provides a unified treatment for models that contain both predictor variables measured without error and predictor variables measured with error.

Proof By the assumptions of the model, MLE are given by (3) provided that $\hat{\Sigma}_{\pi\pi}$ is positive definite and $\hat{\Sigma}_{ww} = \begin{pmatrix} \hat{\sigma}_{vv} & \Sigma_{uv}' \\ \Sigma_{uv} & \Sigma_{uu} \end{pmatrix}$ is at least p.s.d. Note that

$$\begin{aligned}
 |\hat{\Sigma}_{uv}| &= \left| \begin{pmatrix} 1 & \mathbf{b}' \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \hat{\Sigma}_{uv} \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{b} & \mathbf{I} \end{pmatrix} \right| \\
 &= \left| \begin{matrix} \hat{\sigma}_{vv} + \mathbf{b}' \Sigma_{uv} + \Sigma_{uv}' \mathbf{b} + \mathbf{b}' \Sigma_{uu} \mathbf{b} & \Sigma_{uv}' + \mathbf{b}' \Sigma_{uu} \\ \Sigma_{uv}' + \Sigma_{uu} \mathbf{b} & \Sigma_{uu} \end{matrix} \right|, \text{ for any } \mathbf{b}.
 \end{aligned}$$

Now, from the theory of generalized inverse,

$$\Sigma_{uv}' + \mathbf{b}' \Sigma_{uu} = \mathbf{0}'$$

has a solution for \mathbf{b} iff $\mathbf{b}' = -\Sigma_{uv}' \Sigma_{uu}^-$, where Σ_{uu}^- is a generalized inverse of Σ_{uu} . Thus, $\mathbf{b}' = -\Sigma_{uv}' \Sigma_{uu}^-$ is a solution to the equation, $\Sigma_{uv}' + \mathbf{b}' \Sigma_{uu} = \mathbf{0}'$. Since Σ_{uu}^- can be any generalized inverse, choose $\Sigma_{uu}^- = \Sigma_{uu}^+$, the Moore-Penrose generalized inverse. Then $|\hat{\Sigma}_{uv}| = (\hat{\sigma}_{vv} - \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}) |\Sigma_{uu}|$, so $|\hat{\Sigma}_{uv}|$ is at least p.s.d. iff Σ_{uu} is at least p.s.d. (true by assumption), and $\hat{\sigma}_{vv} \geq \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}$. ■

Thus, if Σ_{uu} and Σ_{uv} are known and (3) results in ML estimators for which $\hat{\sigma}_{vv} \leq \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}$, then one cannot set $\hat{\sigma}_{vv} = 0$. The boundary value is $\hat{\sigma}_{vv} = \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}$.

Estimators in (3) should meet some conditions given in Theorem 2.1 to be MLE of linear structural MEM. There is a convenient method to check whether these conditions are satisfied, and it is given in Theorem 2.2. It is composed of three parts. Part i) describes sufficient and necessary condition for the estimators in (3) to be MLE. Parts ii) and iii) mention on MLE of the model parameters when the condition given in Part i) is not satisfied. Proofs of Part ii) and iii) are omitted.

Theorem 2.2 Let $\hat{\lambda}_{l+1}^{-1} \leq \hat{\lambda}_{l+2}^{-1} \leq \dots \leq \hat{\lambda}_k^{-1} \leq \hat{\lambda}_{k+1}^{-1}$ be the positive values of λ^{-1} that satisfy

$$|\bar{\Sigma}_{aa} - \lambda^{-1} \mathbf{m}_{zz}| = 0, \tag{4}$$

where

$$\bar{\Sigma}_{aa} = \begin{pmatrix} \bar{\sigma}_{aa} & \Sigma_{uv}' \\ \Sigma_{uv} & \Sigma_{uu} \end{pmatrix}, \text{ and } \bar{\sigma}_{aa} = \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}.$$

Then,

i) $\hat{\Sigma}_{\pi\pi} = \mathbf{m}_{xx} - \Sigma_{uu}$ is p.d. and

$$\hat{\sigma}_{vv} = \mathbf{m}_{yy} - (\mathbf{m}_{xy} - \Sigma_{uv})' (\mathbf{m}_{xx} - \Sigma_{uu})^{-1} (\mathbf{m}_{xy} - \Sigma_{uv}) > \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}, \text{ iff } \hat{\lambda}_{k+1} > 1.$$

That is, estimators given in (3) are MLE of model (2) iff $\lambda_{k+1} > 1$.

ii) If $\lambda_{k+1} < 1$ and $\lambda > 1$, MLE of β_1 and $\Sigma_{\pi\pi}$ are


$$\hat{\beta}_1 = (\mathbf{m}_{xx} - \lambda_{k+1} \Sigma_{uu})^{-1} (\mathbf{m}_{xy} - \lambda_{k+1} \Sigma_{uv}), \hat{\Sigma}_{\pi\pi} = \mathbf{m}_{xx} - \lambda_{k+1} \Sigma_{uu},$$

and $\hat{\sigma}_{uv} = \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}$.

iii) If $\lambda_k \leq 1$, MLE of $\Sigma_{\pi\pi}$ is singular and the estimator of β_1 is indeterminate.



i) "⇒"

 When $\Sigma_{uv} = 0$.

Since $\Sigma_{uv} = 0$, we have $\tilde{\Sigma}_{aa} = \begin{pmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Sigma_{uu} \end{pmatrix}$. Therefore, (4) becomes

$$\left| \begin{pmatrix} m_{yy} & \mathbf{m}'_{xy} \\ \mathbf{m}_{xy} & \mathbf{m}_{xx} - \Sigma_{uu} \end{pmatrix} - (\lambda - 1) \begin{pmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Sigma_{uu} \end{pmatrix} \right| = 0.$$

From the given conditions, we have

$$\mathbf{m}_{xx} - \Sigma_{uu} > \Phi \text{ and } m_{yy} - \mathbf{m}'_{xy} (\mathbf{m}_{xx} - \Sigma_{uu})^{-1} \mathbf{m}_{xy} > 0,$$

ensuring that


$$\begin{pmatrix} m_{yy} & \mathbf{m}'_{xy} \\ \mathbf{m}_{xy} & \mathbf{m}_{xx} - \Sigma_{uu} \end{pmatrix} > \Phi$$

and, therefore, that $\lambda \neq 1$. So, the above determinantal equation becomes

$$\left| \begin{pmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Sigma_{uu} \end{pmatrix} - (\lambda - 1)^{-1} \begin{pmatrix} m_{yy} & \mathbf{m}'_{xy} \\ \mathbf{m}_{xy} & \mathbf{m}_{xx} - \Sigma_{uu} \end{pmatrix} \right| = 0,$$

which implies $(\lambda - 1)^{-1} \geq 0$ since $\begin{pmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Sigma_{uu} \end{pmatrix}$ is p.s.d. Therefore, we get the final result,

$$\lambda_{k+1} > 1.$$

 When $\Sigma_{uv} \neq 0$.

Consider $\left| \begin{pmatrix} m_{yy} & \mathbf{m}'_{xy} \\ \mathbf{m}_{xy} & \mathbf{m}_{xx} \end{pmatrix} - \lambda \begin{pmatrix} \Sigma_{uv}' \Sigma_{uv}^+ \Sigma_{uv} & \Sigma_{uv}' \\ \Sigma_{uv} & \Sigma_{uu} \end{pmatrix} \right| = 0$, which is equivalent to

$$\left| \begin{pmatrix} m_{yy} - \Sigma_{uv}' \Sigma_{uv}^+ \Sigma_{uv} & \mathbf{m}'_{xy} - \Sigma_{uv}' \\ \mathbf{m}_{xy} - \Sigma_{uv} & \mathbf{m}_{xx} - \Sigma_{uu} \end{pmatrix} - (\lambda - 1) \tilde{\Sigma}_{aa} \right| = 0.$$

From the given conditions, we have

$$m_{xx} - \Sigma_{uv} > \Phi, \text{ and}$$

$$m_{yy} - \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv} - (m_{xy} - \Sigma_{uv})' (m_{xx} - \Sigma_{uv})^{-1} (m_{xy} - \Sigma_{uv}) > 0,$$

ensuring that

$$\begin{pmatrix} m_{yy} - \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv} & m_{xy} - \Sigma_{uv}' \\ m_{xy} - \Sigma_{uv}' & m_{xx} - \Sigma_{uv} \end{pmatrix} > \Phi.$$

Following the remaining steps of Case 1 reveals again that $\hat{\lambda}_{k+1} > 1$.

“ \Leftarrow ” (Proof only for the Case 2 is provided.)

We know that $\hat{\lambda}_{k+1}$ is the smallest root of (4), and $\hat{\lambda}_{k+1} > 1$ from the given condition. Also, since m_{zz} is p.d., λ_j^{-1} are the eigenvalues of $(m_{zz})^{-1/2} \tilde{\Sigma}_{aa} (m_{zz})^{-1/2}$ from (4). Now let d_j be the j^{th} eigenvalue of $(m_{zz})^{-1} \tilde{\Sigma}_{aa}$. Then $d_j = \lambda_{k+2-j}^{-1}$, and $0 \leq d_j < 1$ by the given condition. Utilizing Restlt 2.1,

$$m_{zz} - \tilde{\Sigma}_{aa} = (Q')^{-1} (I - D) Q^{-1} > \Phi,$$

where $D = \text{diag}(d_1, d_2, \dots, d_{k+1})$, since $1 - d_j > 0$. Consequently,

$$\hat{\Sigma}_{\pi\pi} = m_{xx} - \Sigma_{uu} > \Phi, \text{ and } \hat{\sigma}_{vv} > \Sigma_{uv}' \Sigma_{uu}^+ \Sigma_{uv}.$$

ii) & iii) : A proof has been given by Amemiya & Fuller [1] and Anderson, Anderson & Olkin [2] ■

In part ii) of Theorem 2.2, MLE's of β_1 and $\Sigma_{\pi\pi}$ were defined as

$$\hat{\beta}_1 = (m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu})^{-1} (m_{xy} - \hat{\lambda}_{k+1} \Sigma_{uv}), \hat{\Sigma}_{\pi\pi} = m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu}.$$

There is a result that guarantees the positive definiteness of $m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu}$, and it is stated in Lemma 2.1 Σ_{uv} is assumed to be p.d. for the convenience of proof. Similar proof can be applied when Σ_{uv} is p.s.d.

Let the model (2) hold with $\Sigma_{uv} > \Phi$. And let $\hat{\lambda}_{l+1}^{-1} \leq \hat{\lambda}_{l+2}^{-1} \leq \dots \leq \hat{\lambda}_k^{-1} \leq \hat{\lambda}_{k+1}^{-1}$ be the positive values of λ^{-1} that satisfy $|\Sigma_{uv} - \lambda^{-1} m_{zz}| = 0$. (In other words, $\hat{\lambda}_{k+1}$ is the smallest root of $|m_{zz} - \lambda \Sigma_{uv}| = 0$) Then, with probability one,

i) $m_{zz} - \hat{\lambda}_{k+1} \Sigma_{uv}$ has rank k ,

ii) $m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu}$ is nonsingular.

From $|\Sigma_{ww} - \hat{\lambda}^{-1} m_{zz}| = 0$, we have

$$|(\Sigma_{ww})^{-1/2} m_{zz} (\Sigma_{ww})^{-1/2} - \hat{\lambda}_{k+1} \mathbf{I}| = 0.$$

Let $p_i = (n-1)^{-1/2} (\Sigma_{ww})^{-1/2} (z_i - \bar{z})$, and $P = (p_1 \ p_2 \ \dots \ p_n)$. Then, $(\Sigma_{ww})^{-1/2} m_{zz} (\Sigma_{ww})^{-1/2} = P P'$. By Okamoto [5]'s result, $P P'$ has rank $k+1$ and nonzero eigenvalues of $P P'$ are distinct. Consequently, by Result 2.1,

$$\begin{aligned} m_{zz} - \hat{\lambda}_{k+1} \Sigma_{ww} &= (T')^{-1} (\Lambda - \hat{\lambda}_{k+1} \mathbf{I}) T^{-1} \\ &= (T')^{-1} \text{diag}(\lambda_1 - \hat{\lambda}_{k+1}, \lambda_2 - \hat{\lambda}_{k+1}, \dots, \lambda_k - \hat{\lambda}_{k+1}, 0) T^{-1}, \end{aligned}$$

so that i) is proven. In addition, since $(m_{zz} - \hat{\lambda}_{k+1} \Sigma_{ww}) \begin{pmatrix} 1 \\ -\hat{\beta}_1 \end{pmatrix} = \mathbf{0}$ and rank

$(m_{zz} - \hat{\lambda}_{k+1} \Sigma_{ww}) = k$, we have

$$\begin{pmatrix} 0 & c' \end{pmatrix} (m_{zz} - \hat{\lambda}_{k+1} \Sigma_{ww}) \begin{pmatrix} 0 \\ c \end{pmatrix} = c' (m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu}) c \neq 0, \quad \forall c \neq 0.$$

Therefore $m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu} > \mathbf{0}$ with probability one, and rank $(m_{xx} - \hat{\lambda}_{k+1} \Sigma_{uu}) = k$ ■

Boundary problems associated with ML estimation were discussed and a method was introduced which could be used to check the existence of such problems. One simple example is given below with linear structural MEM with single predictor variable.

Consider the model (2) with $k = 1$. As was mentioned in the first part of this section, σ_{vv} is assumed to be unknown and σ_{uu} is assumed to be known. Further, let's assume that $\sigma_{uv} = 0$ for simple derivation. Then, the positive solution $\hat{\lambda}^{-1}$ of (4) is

$$\hat{\lambda}_2^{-1} = \frac{\sigma_{uu} m_{yy}}{m_{xx} m_{yy} - m_{xy}^2}.$$

Therefore, we have $\hat{\lambda}_1 = \infty$ and $\hat{\lambda}_2 = \frac{m_{xx} m_{yy} - m_{xy}^2}{\sigma_{uu} m_{yy}}$.

Now from Theorem 2.2, we have the following results which are the same as those of Birch [3]

i) $\hat{\sigma}_{\pi\pi} = m_{xx} - \sigma_{uu} > 0$, $\hat{\sigma}_{vv} = m_{yy} - (m_{xx} - \sigma_{uu})^{-1} m_{xy}^2 > 0$ and

$\hat{\beta}_1 = (m_{xx} - \sigma_{uu})^{-1} m_{xy}$ iff $\lambda_2 > 1$, which is equivalent to $m_{yy} m_{xx} - m_{xy}^2 > \sigma_{uu} m_{yy}$.

ii) Suppose $\lambda_2 > 1$, that is $m_{yy} (m_{xx} - \sigma_{uu}) > m_{xy}^2$. Then, by Theorem 2.1 and Part ii) of Theorem 2.2, $\hat{\sigma}_{\pi\pi} = m_{yy}^{-1} m_{xy}^2$, $\hat{\sigma}_{vv} = 0$ and $\hat{\beta}_1 = m_{xy}^{-1} m_{yy}$. ■

◆ REFERENCES ◆

1. Amemiya, Y. and Fuller, W. A., "Estimation for the Multivariate Errors-in-Variables Model with Estimated Error Covariance Matrix," *Annals of Statistics*, 12, 1984, pp. 497~509.
2. Anderson, B. M., Anderson, T. W., and Olkin, I., "Maximum Likelihood Estimators and Likelihood Ratio Criteria in Multivariate Components of Variance," *Annals of Statistics*, 14, 1986, pp. 405~417.
3. Birch, M. W., "A Note on the Maximum Likelihood Estimation of a Linear Structural Relationship," *Journal of the American Statistical Association*, 59, 1964, pp. 1175~1178.
4. Fuller, W. A., *Measurement Error Models*, 1987, New York: Wiley and Sons, Inc..
5. Okamoto, M., "Distinctness of the Eigenvalues of a Quadratic Form in a Multivariate Sample," *Annals of Statistics*, 1, 1973, pp. 763~765.