〈Research Paper〉

Journal of the Korean
Statistical Society
Vol. 13, No. 2, 1984

Average Mean Square Error of Prediction for a Multiple Functional Relationship Model

Bong Jin Yum*

ABSTRACT

In a linear regression model the independent variables are frequently subject to measurement errors. For this case, the problem of estimating unknown parameters has been extensively discussed in the literature while very few has been concerned with the effect of measurement errors on prediction. This paper investigates the behavior of the predicted values of the dependent variable in terms of the average mean square error of prediction(AMSEP). AMSEP may be used as a criterion for selecting an appropriate estimation method, for designing an estimation experiment, and for developing cost-effective future sampling schemes.

1. Introduction

Consider the following relationship among variables u_1, u_2, ..., u_p and v.

$$v=\beta_0+\beta_1 u_1+\beta_2 u_2+\cdots+\beta_p u_p \tag{1}$$

where β's are unknown constants. In an experiment to estimate these unknowns, suppose one cannot measure u's and v exactly, but only observes

$$\left.\begin{aligned} x_{ij}&=u_{ij}+\delta_{ij} \\ y_i&=v_i+\varepsilon_i \end{aligned}\right\} \quad \begin{aligned} i&=1, 2, \cdots, n \\ j&=1, 2, \cdots, p \end{aligned} \tag{2}$$

where δ_{ij} and ε_i represent random measurement errors. The model described in Eqs. (1) and (2) is frequently called a (multiple) errors-in-variables model (EVM) in the literature. The EVM is further classified into functional and structural model if the variables in $Eq.$ (1) is fixed and random, respectively (Kendall and Stuart, 1979).

The problem of estimating β's in $Eq.$ (1) has been extensively discussed, especially

* Korea Advanced Institute of Science and Technology, P.O. Box 150, Chongyangni, Seoul, Korea

for the simple EVM (e.g., see Madansky (1959), Moran (1971), Kendall and Stuart (1979)). However, little work has been done on the prediction problem in the EVM context. For a multiple structural model Lindley (1947) established a condition under which a linear regression of y on $x_1, x_2, \ldots, x_p$ exists so that a most likely value of future y may be predicted from future observed $x_1, x_2, \ldots, x_p$. Ganse et al. (1983) extended Lindley's work to the case where the estimation and prediction population may be different. Recently, Yum and Neuhardt (1984) considered a prediction problem for a simple functional relationship model and compared the relative performance of ordinary and grouping least squares method using the average mean square error of prediction ($AMSEP$) as a criterion.

The purpose of this paper is to extend the Yum and Neuhardt result by developing $AMSEP$ for a multiple functional relationship model. The following development is general in the sense that it can be applied to any estimation method that generates an estimator with finite mean and variance. Thus $AMSEP$ can be used as a criterion for a discrimination of competing estimators for a multiple functional relationship model. Further, $AMSEP$ may provide an experimenter with useful guidelines for designing an estimation experiment as well as for determining proper level of sampling efforts in the future.

2. Notation and Assumptions

Define

$$\left.\begin{array}{ll} v=(v_1, v_2, \cdots, v_n)' & \\ u_i=(1, u_{i1}, \cdots, u_{ip})' & i=1, 2, \cdots, n \\ y=(y_1, y_2, \cdots, y_n)' & \\ x_i=(1, x_{i1}, \cdots, x_{ip})' & i=1, 2, \cdots, n \\ \beta=(\beta_0, \beta_1, \cdots, \beta_p)' & \\ \varepsilon=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)' & \\ \delta_i=(0, \delta_{i1}, \cdots, \delta_{ip})' & i=1, 2, \cdots, n \\ U'=[u_1, u_2, \cdots, u_n] & \\ X'=[x_1, x_2, \cdots, x_n] & \\ \Delta'=[\delta_1, \delta_2, \cdots, \delta_n] & \end{array}\right\} \quad (3)$$

Then, Eqs. (1) and (2) may be rewritten as

$$\left.\begin{array}{l} X=U+\Delta \\ y=v+\varepsilon \\ \quad =U\beta+\varepsilon \end{array}\right\} \quad (4)$$

Consider an error vector

$$e_i=\begin{bmatrix}\varepsilon_i\\ \delta_i\end{bmatrix},\quad i=1,\ 2,\ \cdots,\ n. \tag{5}$$

We assume that e_i is distributed with mean vector 0 and covariance matrix

$$\begin{bmatrix}\sigma_\varepsilon^2 & 0\\ 0 & \Sigma\end{bmatrix} \tag{6}$$

where $\Sigma=\mathrm{cov}(\delta_i)=\mathrm{diag}(0,\ \sigma_1^2,\ \sigma_2^2,\ \cdots,\ \sigma_p^2)$. We further assume that e_i's are pairwise independent.

3. A Prediction Model

Suppose for the model in *Eq.* (4) an estimate b for β is available. For instance, ordinary least squares estimation yields

$$b=(X'X)^{-1}X'y. \tag{7}$$

Then, using the estimated relationship one may wish to predict v at an unknown $u\in R$, where R is a known region of interest. Again, it is assumed that the future u cannot be measured precisely, but one observes a sequence of replicated observations

$$x_k=u+\delta_k,\quad k=1,\ 2,\ \cdots,\ m \tag{8}$$

where $\delta_k=(0,\ \delta_{k1},\ \cdots,\ \delta_{kp})'$ and δ_{ki}'s are independent with mean 0 and variance σ_i^2. Then, an estimate of v is obtained as

$$\hat{v}=\bar{x}'b \tag{9}$$

where

$$\bar{x}=\frac{1}{m}\sum_{k=1}^{m}x_k=u+\bar{\delta}. \tag{10}$$

In the following section, the property of $\hat{v}$ over R is examined in terms of $AMSEP$.

4. Average Mean Square Error of Prediction

Assume that each element of u is scaled such that the region of interest $R=\{u:-1\leq u_i\leq 1,\ i=1,\ 2,\ \cdots,\ p\}$. $AMSEP$ is then defined as

$$AMSEP=2^{-p}\int_R MSE(\hat{v})du. \tag{11}$$

By definition, the mean square error (MSE) of $\hat{v}$ is given by

$$MSE(\hat{v})=E(\hat{v}-u'\beta)^2$$

$$=\mathrm{Var}(\hat{v})+\{E(\hat{v})-u'\beta\}^2. \tag{12}$$

From *Eq.* (9) and (10)

$$E(\hat{v})=E(\bar{x}'b)=E\{(u+\bar{\delta})'b\}=u'E(b) \tag{13}$$

assuming that the future measurement errors and b are independent. If we define the bias in b as

$$\gamma=E(b)-\beta \tag{14}$$

then,

$$MSE(\hat{v})=\mathrm{Var}(\hat{v})+(u'\gamma)^2. \tag{15}$$

However,

$$\begin{aligned}\mathrm{Var}(\hat{v})&=E(\hat{v}^2)-\{E(\hat{v})\}^2\\&=E\{(u+\bar{\delta})'b\}^2-\{u'E(b)\}^2.\end{aligned} \tag{16}$$

To evaluated $E\{(u+\bar{\delta})'b\}^2$, the following theorem is used (a proof is in Appendix).

Theorem. Let f and g be q-variate, independent random vectors distributed with mean vector and covariance matrix $(\bar{f}, F)$and $(\bar{g}, G)$, respectively.

Then,

$$E(f'g)^2=\mathrm{tr}FG+\bar{g}'F\bar{g}+\bar{f}'G\bar{f}-(\bar{f}'\bar{g})^2. \tag{17}$$

Applying the theorem to $E\{(u+\bar{\delta})'b\}^2$ with $E(u+\bar{\delta})=u$ and $\mathrm{Var}(u+\bar{\delta})=\frac{1}{m}\Sigma$, and combining Eqs. (15) and (16), we obtain

$$\begin{aligned}MSE(\hat{v})&=\frac{1}{m}\mathrm{tr}\,\Sigma V+\frac{1}{m}E(b)'\Sigma E(b)+u'Vu+(u'\gamma)^2\\&=\frac{1}{m}\mathrm{tr}\,\Sigma V+\frac{1}{m}(\beta+\gamma)'\Sigma(\beta+\gamma)+u'Vu+u'(\gamma\gamma')u.\end{aligned} \tag{18}$$

where $V=\mathrm{Var}(b)$.

Integrating *Eq.* (18) according to (11) yields

$$\begin{aligned}AMSEP=&\frac{1}{m}\sum_{i=1}^{p}\sigma_i^2\,\mathrm{Var}(b_i)+\frac{1}{m}\sum_{i=1}^{p}(\beta_i+\gamma_i)^2\sigma_i^2\\&+MSE(b_0)+\frac{1}{3}\sum_{i=1}^{p}MSE(b_i).\end{aligned} \tag{19}$$

A detailed derivation of *Eq.* (19) is found in Appendix. Note that if there is no error in u_i, then

$$AMSEP=\mathrm{Var}(b_0)+\frac{1}{3}\sum_{i=1}^{p}\mathrm{Var}(b_i). \tag{20}$$

5. Discussions

Eq. (19) suggests several strategies we can adopt to reduce $AMSEP$. First, the future sample size m may be increased. In this case however, unnecessarily large m should be avoided since it may be costly with little additional benefit to the reduction of $AMSEP$(as can be seen in *Eq.*(19), $AMSEP$ contains terms which are not reduced by increasing m). Secondly, we may reduce $AMSEP$ by selecting an appropriate estimation method. In this sense, the $AMSEP$ may serve as a criterion for a discrimination of competing estimation methods.

Considering the above Yum and Neuhardt (1984) provided a detailed analysis of the prediction problem for a simple functional relationship model when replicated observations are available. Among others they found that increasing n does not necessarily reduce $AMSEP$, and in most cases m needs not be increased beyond 6 or 8.

In case an estimate of $AMSEP$ is desired for a given estimation method, several unknowns in *Eq.* (19) must be estimated. Usually exact expression for V and γ are very complicated or difficult to obtain, and therefore, some approximation is necessary. For instance, if ordinary least squares estimation is used, Davies and Hutton (1975) show that for large n,

$$V \simeq \frac{1}{n}\left\{\left(\frac{1}{n}U'U+\Sigma\right)^{-1}(\sigma_\varepsilon^2+\beta'\Sigma\beta)\right\} \tag{21}$$

$$\gamma \simeq \left(\frac{1}{n}U'U+\Sigma\right)^{-1}\Sigma\beta. \tag{22}$$

In addition, Seber (1977) shows that

$$E(X'X)=\left(\frac{1}{n}U'U+\Sigma\right). \tag{23}$$

Then, a "rough" estimate of $AMSEP$ may be obtained by using $X'X$ for $\left(\frac{1}{n}U'U+\Sigma\right)$, b for β, and the usual error mean square for σ_ε^2. An estimate of Σ may be obtained from the historical data or from seperate experiments.

REFERENCES

(1) Davies, R.B. and Hutton, B. (1975) The Effects of Errors in the Independent Variables in Linear Regression, Biometrika, Vol. 62, 383-391.

(2) Ganse, R.A., Amemiya, Y., and Fuller, W.A. (1983) Prediction When Both Variables are

Subject to Error, with Application to Earthquake Magnitudes, J. Amer. Statist. Ass., Vol. 78, 761-765.

(3) Kendall, M. and Stuart, A. (1979) The Advanced Theory of Statistics (Fourth ed.), Vol. 2, Hafner, New York.

(4) Lindley, D.V. (1947) Regression Lines and the Linear Functional Relationship, J. Roy. Statist. Soc., Vol. 9, 218-244.

(5) Madansky, A. (1959) The Fitting of Straight Lines When Both Variables are Subject to Error, J. Amer. Statist. Ass., Vol. 54, 173-205.

(6) Moran, P.A.P. (1971) Estimating Structural and Functional Relationships, J. Multivariate Analysis, Vol. 1, 232-255.

(7) Yum, B.J. and Neuhardt, J.B. (1984) Analysis of the Prediction Problem in a Simple Functional Relationship Model, *IIE* Transactions, Vol. 16, 177-184.

APPENDIX

1. Proof of the Theorem

Let $F=(f_{ij})$ and $G=(g_{ij})$. Then

$$E(f'g)^2=E\{\sum_{i=1}^{q}\sum_{j=1}^{q}f_i g_i f_j g_j\}$$

$$=\sum_{i=1}^{q}\sum_{j=1}^{q}E(f_i f_j)\ E(g_i g_j)$$

$$=\sum_{i=1}^{q}\sum_{j=1}^{q}(f_{ij}+\bar{f}_i\bar{f}_j)(g_{ij}+\bar{g}_i\bar{g}_j)$$

$$=\sum_{i=1}^{q}\sum_{j=1}^{q}f_{ij}g_{ij}+\sum_{i=1}^{q}\sum_{j=1}^{q}f_{ij}\bar{g}_i\bar{g}_j$$

$$+\sum_{i=1}^{q}\sum_{j=1}^{q}g_{ij}\bar{f}_i\bar{f}_j+\sum_{i=1}^{q}\sum_{j=1}^{q}\bar{f}_i\bar{f}_j\bar{g}_i\bar{g}_j$$

$$=\mathrm{tr}\,FG+\bar{g}'F\bar{g}+\bar{f}'G\bar{f}+(\bar{f}'\bar{g})^2.$$

2. Derivation of AMSEP

The elements of u, γ, Σ, and V are partitioned as follows.

$$u=\begin{pmatrix}1\\ \tilde{u}\end{pmatrix}_{p}^{1},\qquad \gamma=\begin{pmatrix}\gamma_0\\ \tilde{\gamma}\end{pmatrix}_{p}^{1}$$

$$\Sigma=\begin{bmatrix}0 & 0\\ 1 & \tilde{\Sigma}\end{bmatrix}_{p}^{1},\qquad V=\begin{bmatrix}v_{00} & v_{01}'\\ v_{01} & \tilde{V}\end{bmatrix}_{p}^{1}$$

$$1\quad p\qquad\qquad 1\quad p$$

Inserting the above into *Eq.* (18) yields.

$$MSE(\hat{v})=\frac{1}{m}\text{tr}\ \Sigma V+\frac{1}{m}(\beta+\gamma)'\Sigma(\beta+\gamma)$$
$$+v_{00}+2\,\tilde{u}'\,v_{01}+\tilde{u}'\,\tilde{V}\tilde{u}$$
$$+\gamma_0{}^2+2\,\gamma_0\tilde{r}'\,\tilde{u}+\tilde{u}'(\tilde{r}\tilde{r}')\,\tilde{u}.$$

Integrating the above according to *Eq.* (11) yields

$$AMSEP=\frac{1}{m}\text{tr}\ \Sigma V+\frac{1}{m}(\beta+\gamma)'\ \Sigma(\beta+\gamma)+v_{00}+\gamma_0{}^2$$
$$+2^{-p}\int_{\tilde{u}}2(\tilde{u}'\,v_{01}+\gamma_0\,\tilde{r}'\,\tilde{u})+\tilde{u}'(\tilde{r}\tilde{r}'+\tilde{V})\,\tilde{u}\,d\tilde{u}$$
$$=\frac{1}{m}\text{tr}\ \Sigma V+\frac{1}{m}(\beta+\gamma)'\Sigma(\beta+\gamma)$$
$$+(v_{00}+\gamma_0{}^2)+\frac{1}{3}\text{tr}(\tilde{r}\tilde{r}'+\tilde{V})$$
$$=Eq.\ (19).$$