

5-1-2004

Estimation Of Multiple Linear Functional Relationships

Amjad D. Al-Nasser

Yarmouk University, Jordan, amjadn@yu.edu.jo

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Al-Nasser, Amjad D. (2004) "Estimation Of Multiple Linear Functional Relationships," *Journal of Modern Applied Statistical Methods*: Vol. 3 : Iss. 1 , Article 22.

DOI: 10.22237/jmasm/1083370920

Estimation Of Multiple Linear Functional Relationships

Amjad D. Al-Nasser
 Department of Statistics
 Yarmouk University, Jordan

This article deals with multiple linear functional relationships models. Two robust estimations procedure are proposed to estimate the model, based on Generalized Maximum Entropy and Partial Least Square. They are distribution free and do not rely (so much) on classical assumptions. The experiments showed that the GME approach outperforms the PLS in terms of mean squares of errors (MSE). Empirical examples are studied.

Key words: Multiple linear functional relationships, generalized maximum entropy, partial least square

Introduction

Consider a set of mathematical variables $(\xi_1, \xi_2, \dots, \xi_K)$ that connected by the relation $\eta = \alpha + \beta_1 \xi_1 + \beta_2 \xi_2 + \dots + \beta_K \xi_K$ and the observed values

$$(x_{ik}, y_i), i = 1, 2, \dots, n, k = 1, 2, \dots, K, n > K > 1,$$

are such that

$$\begin{aligned} y_i &= \eta_i + \varepsilon_i \\ x_{ik} &= \xi_{ik} + \delta_{ik} \end{aligned} \quad (1)$$

where

$$\eta_i = \alpha + \beta_1 \xi_{i1} + \beta_2 \xi_{i2} + \dots + \beta_K \xi_{iK}$$

which can be defined as MLFR, where α is the intercept, β is the slope vector $K \times 1$, ξ unknown, unobservable latent variable with dimension $K \times n$, ε and δ , are mutually independent unobservable disturbances terms with zero means and finite variances. They are assumed to be independent of (ξ_{ik}, η_i) .

Amjad D. Al-Nasser: Department of Statistics, Faculty of Science. Yarmouk University, Irbid Jordan. Email him at amjadn@yu.edu.jo.

Methodology

This article considers two estimation approaches. One is the PLS developed by Wold (1975). The algorithm is presented in Helland (1990) and Geladi (1986). The second is a theoretically attractive estimation approach based on Shannon's (1948) entropy, which is called the Generalised Maximum Entropy. It was developed by Golan et al. (1996) for limited data. Subsequently, Al-Nasser et al. (2000), Abdullah (2000), and Al-Nasser (2003) used it for estimating Errors-in-Variables models. The GME method is a nonlinear programming approach to determine the most random probability distribution subject to the given information (see (1)). Without loss of generality, by assuming the intercept is equal to zero, the model can be rewritten as

$$y_i = \sum_{k=1}^K \beta_k (x_{ik} - \delta_{ik}) + \varepsilon_i, i = 1, 2, \dots, n \quad (2)$$

Because the unknowns are not in probability distribution form, their possible outcomes values are reparametrized as convex combination presented as expected value of discrete random variable:

$$\beta_k = \sum_{h=1}^H z_{hk} p_{hk}, \sum_{h=1}^H p_{hk} = 1, k = 1, 2, \dots, K, H \geq 2 \quad (3)$$

The restriction imposed on the parameter space through z reflects à priori knowledge about these parameters. However, if the possible values of the parameters are known from the theory, then z is specified accordingly. If not, then z is specified to be uniformly symmetric around 0 with high - low upper bounds. For example, $z = (-c, 0, c)$, c large value.

For the simulation study here, assume $z_{hk} = z_h \forall k = 1, 2, \dots, K$. In addition, the unobservable ξ_{ik} can be obtained from the observed data values x_{ik} and the disturbance δ_{ik} can be treated in similar fashion:

$$\delta_{ik} = \sum_{t=1}^T v_{ikt}^* w_{ikt}^* , \sum_{t=1}^T w_{ikt}^* = 1 , k = 1, 2, \dots, K , i = 1, 2, \dots, n, T \geq 2 \quad (4)$$

$$\varepsilon_i = \sum_{j=1}^J v_{ij} w_{ij} , \sum_{j=1}^J w_{ij} = 1 , i = 1, 2, \dots, n, J \geq 2 \quad (5)$$

The actual bounds for v_t^* and v_j depend on the observed sample as well as any conceptual or empirical information about the underlying error. However, if such conceptual or empirical information does not exist, then v_t^* , v_j may be specified to be uniformly and symmetrically distributed around zero. Chebychev's inequality or 3-sigma rule Pukelsheim (1994) may be used as conservative means of specifying sets of error bounds. Under these reparameterizations, the statistical MLFR model (2) can be rewritten as

$$y_i = \sum_{k=1}^K \left[\left\{ x_{ik} - \sum_{t=1}^T v_{ikt}^* w_{ikt}^* \right\} \cdot \sum_{h=1}^H z_{hk} p_{hk} \right] + \sum_{j=1}^J v_{ij} w_{ij} \quad (6)$$

Given the reparameterization inverse problem (6) involving the unknown and the unobservable p, w^*, w , the GME problem as maximizing the dual objective function may be stated in scalar summations with three nonnegative probability components:

$$\begin{aligned} & \text{Maximize } H(p, w^*, w) \\ & = - \sum_{h=1}^H \sum_{k=1}^K p_{hk} \log(p_{hk}) \\ & - \sum_{i=1}^n \sum_{j=1}^J w_{ij} \log(w_{ij}) - \sum_{i=1}^n \sum_{k=1}^K \sum_{t=1}^T w_{ikt}^* \log(w_{ikt}^*) \end{aligned}$$

Subject to the consistency constraints (data)

$$y_i = \sum_{k=1}^K \left[\left\{ x_{ik} - \sum_{t=1}^T v_{ikt}^* w_{ikt}^* \right\} \cdot \sum_{h=1}^H z_{hk} p_{hk} \right] + \sum_{j=1}^J v_{ij} w_{ij} , i = 1, 2, \dots, n$$

and adding-up normalization constraints

- (I) $\sum_{h=1}^H p_{hk} = 1 , k = 1, 2, \dots, K$
- (II) $\sum_{j=1}^J w_{ij} = 1 , i = 1, 2, \dots, n$
- (III) $\sum_{t=1}^T w_{ikt}^* = 1 , i = 1, 2, \dots, n , k = 1, 2, \dots, K$

Solving the optimization problem of the (HK+nJ+nKT) equations yields the optimal and unique solution:

$$\hat{p}_{hk} = \frac{\exp \left[- \sum_{k=1}^K z_{hk} \left\{ \sum_{i=1}^n \hat{\gamma}_i \left(x_{ih} - \sum_{t=1}^T \hat{v}_{iht}^* \hat{w}_{iht}^* \right) \right\} \right]}{\sum_{h=1}^H \exp \left[- \sum_{k=1}^K z_{hk} \left\{ \sum_{i=1}^n \hat{\gamma}_i \left(x_{ih} - \sum_{t=1}^T \hat{v}_{iht}^* \hat{w}_{iht}^* \right) \right\} \right]} \quad (7)$$

$$\hat{w}_{ij} = \frac{\exp[-\hat{\gamma}_i v_{ij}]}{\sum_{i=1}^n \exp[-\hat{\gamma}_i v_{ij}]} \quad (8)$$

$$\hat{w}_{iht}^* = \frac{\exp \left[- \hat{\gamma}_i v_t^* \sum_{k=1}^K z_{hk} \hat{p}_{hk} \right]}{\sum_{i=1}^n \sum_{h=1}^H \exp \left[- \hat{\gamma}_i v_t^* \sum_{k=1}^K z_{hk} \hat{p}_{hk} \right]} \quad (9)$$

where $\hat{\gamma}_i, i = 1, 2, \dots, n$ are Lagrangian multipliers. The estimated parameters can be obtained by substitute (7) in (3);

$$\hat{\beta}_k = \sum_{h=1}^H z_{hk} \hat{p}_{hk} \quad k = 1, 2, \dots, K$$

Similarly, by substitution of (8) and (9) in (4) and (5) respectively, the estimated values of the unobservable variables may be obtained.

Sampling Experiments

Now, consider the performance of the GME and PLS method in estimating the parameters of the MLFR models in (2). Some experiments are carried out to choose the supported weights of the unknown elements under the following conditions:

- (i) Generate 500 samples each of size $n = 15$ observations with $K = 3$ parameters.
- (ii) ξ_{ik} are initialised to be in the interval $\left[-\left\lceil \frac{nK}{2} \right\rceil, \left\lceil \frac{nK}{2} \right\rceil \right]$ with increment 1, where $\lceil * \rceil$ is the greater number less than or equal to $*$.
- (iii) Assume that $\beta_k = 1, \forall k = 1, 2, \dots, K$.
- (iv) The disturbance δ_{ik} and ε_i are generated from Standard Normal distribution.
- (v) Simulate the observed values from the following equations:

$$x_{ik} = \xi_{ik} + \delta_{ki} \quad , i = 1, 2, \dots, n, k = 1, 2, \dots, K$$

$$y_i = \xi_{ik} + \varepsilon_i \quad , i = 1, 2, \dots, n, k = 1, 2, \dots, K .$$

Choice of the index parameters space (Experiment 1)

The aim of this experiment is to chose the disturbance index by varying its values as J, T = 3, 4, 5, 6, 7. The support space of the parameters β_k are fixed to be indexed into 5 values in the interval $[-c, c]$ for $c = 10$. The results of this experiment are tabulated as follows:

Table 1. Choice of residual indices for GME-MLFR model.

Method	MSE($\hat{\beta}_1$)	MSE($\hat{\beta}_2$)	MSE($\hat{\beta}_3$)
PLS	0.0672	0.0659	0.0668
GME (J & T =3)	0.0242	0.0224	0.0203
GME (J & T =4)	0.0243	0.0223	0.0203
GME (J & T =5)	0.0246	0.0222	0.0202
GME (J & T =6)	0.0248	0.0221	0.0201
GME (J & T =7)	0.0250	0.0220	0.0201

In the same way, an experiment is carried out to choose the index value of the parameter space z. The values of z fixed to be located in the interval $[-10, 10]$. The results are tabulated as follows:

Table 2. Choice of parameters index in GME-MLFR model.

Method	MSE($\hat{\beta}_1$)	MSE($\hat{\beta}_2$)	MSE($\hat{\beta}_3$)
PLS	0.0672	0.0659	0.0668
GME (M=3)	0.0249	0.0221	0.0201
GME (M=4)	0.0249	0.0221	0.0201
GME (M=5)	0.0247	0.0222	0.0202
GME (M=6)	0.0245	0.0222	0.0203
GME (M=7)	0.0244	0.0223	0.0203

This experiment suggests that regardless of the index values of parameter supports space the GME is superior than PLS estimates and more robust.

Choice of parameter space values (Experiment 2)

Under the above conditions and by fixing five index values for the parameters in the interval $[-10, 10]$, this experiment is carried out to find the suitable parameter space of the disturbance part. The choice of the supported values was obtained depending on Chebyshev's inequality, and the following table shows the simulated results:

Table 3 Choice of residuals space values for GME-MLFR model.

Method	MSE($\hat{\beta}_1$)	MSE($\hat{\beta}_2$)	MSE($\hat{\beta}_3$)
PLS	0.0672	0.0659	0.0668
GME	0.0246	0.0222	0.0202
{[-3S,3S]}			
GME	0.0241	0.0224	0.0205
{[-4S,4S]}			
GME	0.0237	0.0226	0.0208
{[-5S,5S]}			
GME	0.0235	0.0226	0.0210
{[-6S,6S]}			
GME	0.0233	0.0227	0.0213
{[-7S,7S]}			

Table.4 Choice of parameter space values of GME-MLFR model.

Method	MSE($\hat{\beta}_1$)	MSE($\hat{\beta}_2$)	MSE($\hat{\beta}_3$)
PLS	0.0672	0.0659	0.0668
GME	0.0234	0.0226	0.0211
{z : [-5,5]}			
GME	0.0246	0.0222	0.0202
{z: [-10, 10]}			
GME	0.0288	0.0203	0.0194
{z: [-50, 50]}			
GME	0.0293	0.0201	0.0194
{z: [-100, 100]}			
GME	0.0294	0.0200	0.0193
{z: [-500, 500]}			
GME	0.0294	0.0200	0.0193
{z: [-1000, 1000]}			

This experiment confirms the superiority of the GME over the PLS in all choices of the support values. Note that the 3-sigma choice gives better estimates with respect to both parameters as shown in Table 3.

Increasing the sample size and the number of parameters (Experiment 3)

Consider increasing the sample size $n = 15,20,30,40$, and varying number of parameters, i.e., $K= 2,3,4$. The parameters index is fixed to be three, with support values $[-3S,0,3S]$ for the disturbance parts and $[-10, 10]$ for the slopes. The results were obtained and are displayed in Tables 5 and 6. The tabulated results demonstrate

that the GME approach performs better than the PLS with respect to their MSE for any sample size and any number of unknown parameters.

Empirical Example

The data were collected via interviews with seventeen customers in order to gain a better understanding of service loyalty and identify salient factor affecting its development Al-Nasser (2000). Two types of services, i.e., cellular phone and TV stations, were selected for the study. For this example, the model is

$$\eta_i = \beta_1 \xi_{i1} + \beta_2 \xi_{i2}, i = 1,2,\dots,17$$

such that

$$x_{i1} = \xi_{i1} + \delta_{i1}$$

$$x_{i2} = \xi_{i2} + \delta_{i2}$$

and

$$y_i = \eta_i + \varepsilon_i$$

where η_i = Customer Loyalty, ξ_{i1} = Service Image, and ξ_{i2} = Customer Satisfaction.

This model is solved by using both approaches (GME and PLS). The results are shown in Table 7. In addition, Figure 1 depicts a distinction between the observed error of the customer loyalty by using both methods with the observed errors closely centred around zero. This appears to support that the notion that GME is a more robust alternative than PLS.

TABLE.7 Estimated values of the parameters using the PLS and GME for customer satisfaction data.

Method	$\hat{\beta}_1$	$\hat{\beta}_2$
PLS	0.3523	0.3682
GME	1.8746	1.0876

Table. 5 Performance of the GME-MLFR for various sample sizes and number of parameters.

		N	15	20	30	40
K						
2	MSE($\hat{\beta}_1$)		0.0267	0.0199	0.0105	0.0075
	MSE($\hat{\beta}_2$)		0.0197	0.0144	0.0120	0.0092
3	MSE($\hat{\beta}_1$)		0.0236	0.0157	0.0114	0.0083
	MSE($\hat{\beta}_2$)		0.0226	0.0168	0.0107	0.0080
4	MSE($\hat{\beta}_3$)		0.0208	0.0183	0.0110	0.0086
	MSE($\hat{\beta}_1$)		0.0229	0.0140	0.0113	0.0079
	MSE($\hat{\beta}_2$)		0.0221	0.0162	0.0096	0.0071
	MSE($\hat{\beta}_3$)		0.0232	0.0186	0.0142	0.0116
	MSE($\hat{\beta}_4$)		0.0215	0.0188	0.0102	0.0077

Table.6 Performance of the PLS MLFR for various sample sizes and number of parameters.

		N	15	20	30	40
K						
2	MSE($\hat{\beta}_1$)		0.0651	0.0448	0.0325	0.0244
	MSE($\hat{\beta}_2$)		0.0651	0.0448	0.0326	0.0244
3	MSE($\hat{\beta}_1$)		0.0672	0.0448	0.0333	0.0249
	MSE($\hat{\beta}_2$)		0.0659	0.0449	0.0334	0.0250
4	MSE($\hat{\beta}_3$)		0.0668	0.0449	0.0333	0.0250
	MSE($\hat{\beta}_1$)		0.0651	0.0449	0.0325	0.0244
	MSE($\hat{\beta}_2$)		0.0651	0.0449	0.0236	0.0244
	MSE($\hat{\beta}_3$)		0.0651	0.0449	0.0325	0.0244
	MSE($\hat{\beta}_4$)		0.0651	0.0449	0.0325	0.0244

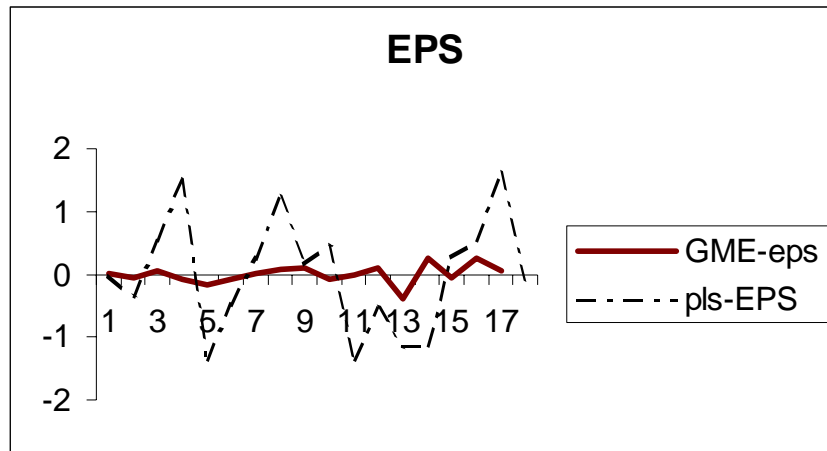


FIGURE 1 Comparison of the observed errors of customer loyalty using GME and PLS.

Conclusion

The results presented in this article provide strong evidence about the robust performance of the GME compared with the traditional methods that have been used to deal with the MLFR. In fact, the PLS that has become the conventional approach in measurement error model fails to outperform the GME in all the cases studied.

From the simulation study, it is evident that with various sample sizes and number of parameters, the MSE of GME estimates are smaller than those of the PLS. Therefore, the GME should provide a better alternative for the MLFR compared with the other methods.

References

- Mokhtar A., Al-Nasser. D. A., & Nooreha, H. (2000). Evaluating functional relationship between image, customer satisfaction, and loyalty using general maximum entropy. *Total Quality Management*, 11, (6), 826-829.
- Geladi, P., & Kowalski, R. B. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, 185, 1-17.
- Golan, A., Judge, G. & Miller, D. (1996). *A maximum entropy econometrics: robust estimation with limited data*. New York: Wiley.
- Helland, S. I. (1990). Partial least squares regression and statistical model. *Scandinavian Journal of Statistics*, 17, 97-114.
- Al-Nasser, D. A., Abdullah, B. N. & Wan Endut, W. J. (2000). On robust estimation of error in variables models by using generalized maximum entropy. International conference on mathematics and its applications in the new millennium. Dept. of Mathematics. Universiti Putra Malaysia. 18-21 July, 279-287.
- Al-Nasser, D. A., (2000). Generalised maximum entropy estimation approach to errors-in-variables models. Unpublished Doctoral dissertation, National University of Malaysia.
- Amjad, D. Al-Nasser, D. A. (2003). Customer satisfaction measurement models: generalized maximum entropy approach. *Pakistan Journal of Statistics*, 19, (2), 213-226.
- Pukelsheim, F. (1994). The three sigma rule. *The American Statistician*, 48, (2), 88-91.
- Shannon C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- Wold, H. (1975). Path models with latent variables: The NIPLAS approach. In H. M. Blalock (Ed). *Quantitative sociology: International perspective on Mathematical and Statistical Modeling*. New York