Journal of the
Military Operations Research
Society of Korea, Vol. 8, No. 1
June, 1982

# Combining Regression Model and Time Series Model to a Set of Autocorrelated Data

Jee, Man Won*

**ABSTRACT**

A procedure is established for combining a regression model and a time series model to fit to a set of autocorrelated data. This procedure is based on an iterative method to compute regression parameter estimates and time series parameter estimates simultaneously. The time series model which is discussed is basically AR(p) model, since MA(q) model or ARMA(p,q) model can be inverted to AR($\infty$) model which can be approximated by AR(p) model. The procedure discussed in this article is applied in general to any combination of regression model and time series model.

## 1. Introduction

A large part of economic data are autocorrelated. Given a set of n observations $\{Y_t;\ t=1,2,3,\ ....,\ n\}$ what we usually proceed is to remove linear effects by introducing a set of explanatory variables When the residuals are found to be autocorrelated (a simple run test or Durbin-Watson test will show this) we attempt to introduce time series analysis to the autocorrelated disturbance terms. In this case both regression model and time series model are combined and thus parameters for both models have to be estimated simultaneously in such a way that the variance of residual terms after fitting both regression and time series models is minimized.

* Korea Institute of Defense Analyses

## 2. Combining Regression Model and Time Series Model

Without loss of generality assume our basic regression model be $Y_t = \alpha + \beta X_t + \varepsilon_t$ where the residual terms $\{\varepsilon_t;\ t=1,2,\ \ldots,\ n\}$ are autocorrelated. For the autocorrelated residuals we would try AR(p) model, MA(q) model or more generally ARMA(p,q) model.

### 2.1 AR(P) model

Start with our basic model $Y_t = \alpha + \beta X_t + \varepsilon_t$ ... (1-1) and $\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_p \varepsilon_{t-p} + a_t$ ... (1-2) where $a_t$ is normally distributed with $N(0,\ \sigma_a^2)$.
Multiplying $\phi_i (i=1,2,\ldots,\ p)$ on both sides of time i lagged data of equation (1-1) [1], we have

$$\phi_1 Y_{t-1} = \phi_1\ (\alpha + \beta X_{t-1} + \varepsilon_{t-1})$$

$$\phi_2 Y_{t-2} = \phi_2\ (\alpha + \beta X_{t-2} + \varepsilon_{t-2})$$

$$\vdots$$

$$\phi_p Y_{t-p} = \phi_p\ (\alpha + \beta X_{t-p} + \varepsilon_{t-p})$$

subtracting from (1-1) the subsequent time lagged equations, we have

$(Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \phi_3 Y_{t-3} - \ldots - \phi_p Y_{t-p}) = \alpha(1 - \phi_1 - \phi_2 - \ldots - \phi_p) + \beta(X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \ldots - \phi_p X_{t-p}) + a_t$ which can be written as $Y_t(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \ldots - \phi_p B^p) = \alpha(1 - \sum_{i=1}^{p} \phi_i) + \beta X_t(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) + a_t$ ... (1-2) in differential form.

Let $V_t = Y_t(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$, $Z_t = X_t(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$ and $\alpha^* = \alpha(1 - \sum_{i=1}^{p} \phi_i)$, then the above differential form becomes $V_t = \alpha^* + \beta Z_t + a_t$ ..... (1-3) where $t = p+1,\ p+2,\ \ldots,\ n$.

Here we are loosing p independent pieces of information in the sample which causes some loss of efficiency.

The loss of efficiency is not significant when p <<< n.

Since $a_t$ is normally distributed and $(X_t - \sum_{i=1}^{p} \phi_i X_{t-i})$ is non-stochastic for t = p+1,..., n, the likelihood function for the parameter of the model $Z_t = \alpha^* + \beta Z_t + a_t$ given

$\{Y_t;\ t=1,2,\ldots, n\}$ is given by $L(\alpha^*, \beta, \phi, \sigma_a^2/Y) =$

$$\frac{1}{(2\pi\sigma_a^2)^{\frac{n-p}{2}}} \cdot \exp\{-\frac{\sum_{t=p}^{n} a_t^2}{2\pi\sigma_a^2}\}$$

where we used $Z_t - \alpha^* - \beta Z_t$ for $a_t$.

$$\text{Log } L = -\frac{n-p}{2}\log 2\pi - \frac{n-p}{2}\log\sigma_a^2 - \frac{1}{2\sigma_a^2}\sum_{t=p}^{n}(Y_t - \phi_i \sum_{i=1}^{p} Y_{t-i}) -$$

$$(\alpha^* + \beta(X_t - \phi_i \sum_{i=1}^{p} X_{t-i})$$

Theoretically we can get MLE(Maximum Likelihood Estimator) for the parameters of transformed model by differentiating log L with respect to $\alpha$, $\beta$, $\phi_1$, $\phi_2$,..., $\phi_p$ and $\sigma_a^2$, but these equations are highly non linear and solutions are not easily tractable.

An alternative way is to choose several sets of values of $\{\phi_1, \phi_2, \ldots, \phi_p\}$ and then choose the set $\{\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p\}$ which gives minimum variance of $a_t$. To do this the procedure would be iterative[2] as follows.

(1) Obtain OLS estimates of original model $Y_t = \alpha + \beta X_t + \varepsilon_t$ and calculate the residuals $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \hat{\varepsilon}_3, \ldots, \hat{\varepsilon}_n$.

(2) Calculate $\hat{\phi}_i = \frac{\sum \varepsilon_t \varepsilon_{t-i}}{\sum \varepsilon_{t-i}^2}$ for i=1,2,3,..., p, and t=i, i+1,..., n.

(3) Construct new variables $V_t$ and $Z_t$ and obtain a new set of OLS estimates of transformed data $V_t = \alpha^* + \beta Z_t + a_t$ using $\hat{\phi}_i$s which are obtained from (2).

(4) Calculate a new set of residuals $\{\hat{\varepsilon}_t,\ t=p, p+1, p+2, \ldots, n\}$ and go to (2) until the values of estimates converge.

Reference [1] showed that final converged values of estimates are identical to the values of the MLE estimates and variance of noise ($\sigma_a^2$) is substantially smaller than the one which we get using only OLS model.

The hypothesis test for $\phi_i$ can be done by either using sample variance $S^2_{\hat{\phi}_i} \approx \frac{1-\hat{\phi}_i^{2}}{n}$ (1) when n is large or by using Durbin Watson test.

## 2.2 MA(q) Model

Theoretically MA(q) process is invertable to infinite number of terms of AR process where parameters of MA(q) process satisfy invertability conditions [3] and the process can be approximated by choosing small number of terms from infinite number of terms of inverted AR process.

For the simplicity consider MA(1) process applied to $\varepsilon_t$.

$$t = a_t - \theta_1 a_{t-1} = a_t(1 - \theta_1 B)$$

we can write $a_t = \frac{t}{1-\theta_1 B} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_1^2\varepsilon_{t-2} + \theta_1^3\varepsilon_{t-3} + \ldots$

As an approximation we can use only p terms; $a_t \approx \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_1^2\varepsilon_{t-2} + .. + \theta_1^p\varepsilon_{t-p}$, where p can be small when $\theta_1 << 1$ and large when $\theta_1$ is close to 1.

When p is large we loose many independent observations which causes material loss of efficiency.

In our model $Y_t = \alpha + \beta X_t + \varepsilon_t$ ..... (2-1)

Multiplying $\theta_1^i$ on both sides of time i lagged data of (2-1)

$$\theta_1 Y_{t-1} = \theta_1(\alpha + \beta X_{t-1} + \varepsilon_{t-1})$$

$$\vdots$$

$$\theta_1^p Y_{t-p} = \theta_1^p (\alpha + \beta X_{t-p} + \varepsilon_{t-p}).$$

Adding above equations together we have

$Y_t(1 + \theta_1 B + \theta_1^2 B^2 + \ldots + \theta_1^p B^p) = \alpha(1 + \theta_1 + \theta_1^2 + \ldots + \theta_1^p) +$
$\beta X_t(1 + \theta_1 B + \theta_1^2 B^2 + \ldots + \theta_1^p B^p) + a_t$

where $a_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-2} + \ldots + \theta_1^p \varepsilon_{t-p}$.

Let $V_t = Y_t(1 + \theta_1 B + \theta_1^2 B^2 + \ldots + \theta_1^p B^p)$, $Z_t = X_t$, $(1 + \theta_1 B + \ldots\ldots + \theta_1^p B^p)$ and $\alpha^* = \alpha(1 + \theta_1 + \theta_1^2 + \ldots + \theta_1^p)$, then our transformed data become;

$V_t = \alpha^* + \beta Z_t + a_t$ to which new set of OLS estimates are established with loss of additional p degrees of freedom in addition to loss of 2 degrees of freedom for the OLS parameters $\alpha$ and $\beta$. This procedure is idetical to AR(p) model.

### 2.3 ARMA(p,q) Model

ARMA(p,q) model can be inverted to pure infinite terms of AR process. We can approximate the process by including a finite number of AR terms combining the two prodecures which are discussed in 2.1 and 2.2.

## 3. Algorithm for Combining regression model and time series model

Given n observations $\{Y_t,\ t=1,2,\ldots,n\}$ and $\{X_{it},\ i=1,2,\ldots,m,\ t=1,2,\ldots,n\}$.

(1) Obtain OLS estimates of $Y_t = \alpha + \beta X_t + \varepsilon_t$, and calculate residuals $\{\hat{\varepsilon}_t;\ t=1,2,\ldots,n\}$

(2) Identify a time series model and compute time series parameter estimates using $\{\hat{\varepsilon}_t,\ t=1,2,\ldots,n\}$

a) if time series model looks pure AR(p) process, then transform the observed data as in 2.1.

b) if time series model looks pure MA(q) process then transform the observed data as in 2.2.

c) if time series model looks ARMA(p,q) process then transform the observed data as in 2.3.

(3) Estimate a new set of regression parameter $(\hat{\alpha}', \hat{\beta}')$ by OLS method from the transformed data $V_t = \Phi(B)Y_t$ and $Z_t = \Phi(B)X_t$ and calculate a new set of residual estimates $\{\hat{\varepsilon}'_t, t=p, p+1, .., n\}$

(4) go to (2) until the values of all parameter estimates (both OLS estimates and time series parameter estimates) converge to a set of numbers.

(5) do diagnostic check including plot of residuals.

(6) use model for forcasting, simulation or systems design.

## 4. Conclusion

The model introduced in this paper is useful to the data that depict a specific phenomenon a part of which can be explained by introducing explanatory variables and the other part by time series model.

The iterative procedure as an algorithm to estimate parameters of both regression model and time series model is basically a generalization and application of the method suggested by Cochrane and Orcutt.

## REFERENCES

1. Jan Kmenta, "Elements of Econometrics", MacMillan, New York, 1971.

2. D. Cochrane and G.H. Orcutt, "Application of Least Squares Regressions to Relationships Containing Autocorrelated Error Terms", Journal of the American Statistical Association, Vol. 44, March 1949, pp. 32-61.

3. George E.P. Box and Gwilym M. Jenkins, "Time Series Analysis", Holden Day, San Francisco, 1970.