[Academic Script]

Autoregressive and Distributed Lag Models

Subject:	Business Economics
Course:	B. A. (Hons.), 5th Semester, Undergraduate
Paper No. & Title:	Paper – 531 Elective Paper Q1 – Advanced Econometrics
Unit No. & Title:	Unit – 3 Time Series Models
Lecture No. & Title:	Lecture – 1 Autoregressive and Distributed Lag Models

Academic Script

1. Introduction

Hello friends nice to meet you. Today we are discussing a very imported topic related to lag variables. You know that when we collect data and apply modelling technique than many times the dependent variable is such that it is always affected by a number of explanatory variables in the current period as well as in the earlier periods. That means all this explanatory variables have accumulative effect upon the dependent variable. This is known as distributed Lag model. We have finite as well as infinite distributed Lag models. Similarly we also have auto regressive model. Both the types of model can be studied by special methods and in this lecture we want to learn about this methods, there limitations, applications etc. in sufficient details so let us start the lecture.

In this category, there are two types of models

(1) Distributed Lag Model

(2) Autoregressive Model

First we illustrate for both these categories before going to complex mathematical expressions.

(a) $Y_t = Constant + 0.6X_t + 0.4X_{t-1} + 0.3X_{t-2} + U_t$

Here Y_t = Consumption in current year t

X_t = Income in the current year t

X_{t-1} = Income in the last year *t-1*

X_{t-2} = Income in the earlier year *t-2*

U_t = Disturbance term

This model is called **Distributed Lag Model (DLM)**

(b) $Y_t = Constant + 0.3X_t + 0.2Y_{t-1} + U_t$

Here Y_t = Consumption in current year t

X_t = Income in the current year t

Y_{t-1} = Consumption in the last year *t-1*

U_t = Disturbance term

This model is called **Autoregressive Lag Model (ALM).**

2. Mathematical form of Distributed Lag Model

Let us write

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + U_t$

Y_t = Consumption in current year t

X_t, X_{t-1}, X_{t-2} are the Incomes in current and successive last 2 years, and U_t is the disturbance term.

Here the total impact on the dependent variable is distributed by successive incomes derived up to the successive last two year's incomes.

We can extend this for <u>finite lag of k periods</u>

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_K X_{t-K} + U_t$

This is called <u>**finite Distributed Lag Model**</u> (FDLM), Coefficient β_0 attached to X_t is called <u>**short run**</u> or <u>**Impact Multiplier.**</u> $\beta_1 \beta_2, \ldots \beta_K$ are called <u>**Delay**</u> or <u>**Interim Multipliers**</u>. $\beta = \sum_{i=0}^{K} \beta_i = \beta_0 + \beta_1 + \beta_2 + \cdots + \beta_K$ is called <u>**Long run**</u> or <u>**Total distributed lag multiplier**</u>. (assuming that β exists). In the case of consumption Income model, β_0 is <u>**short run MPC**</u>, $\beta_1, \beta_2, \ldots \beta_K$ are called <u>**Interim MPC**</u> and β is called <u>**long run MPC**</u>.

In a similar Way, we can define <u>**Infinitely Distributed Lag Model**</u> **(IDLM)** as under

$Y_t = \alpha + (\beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots \text{ up to } \infty) + U_t$

3. Mathematical form of Autoregressive Model

Let us write the model form as

$Y_t = \alpha + \beta X_t + \gamma . Y_{t-1} + U_t$

In this model, the right hand side contains one explanatory variable X in the current year and also the dependent variable Y is the past period (t-1). This model is known as

Autoregressive Model of the first order

4. Some practical illustrations how lags are generated

Few practical illustrations for lag structures are shown below

(1) Consumption in the current year becomes a linear function of the successive streams derived in the subsequent years which follows a DL model.

(2) Demand deposits in banks can be expressed as a linear function of buying successive government Securities in different periods.

(3) The increase in the general price level can be expressed as a linear function of the rate of expansion in money supply in successive periods

(4) Expenditures towards research and development funds can be expressed as a linear function of the increase in productivity during subsequent years.

5. What are the reasons for Lags?

(1) Psychological

Generally the consumption habit of people does not change abruptly following a decrease in price or an increase in income. This in turn helps to generate a Lag.

(2) Technological

Technogical innovations play a key role to generate Lag structure. e.g. In electronics industry there is very rapid development and due to fast advancing technical researches in electronics field day by day it affects its price and consumption pattern reasonably.

(3) Institutional

Due to contractual obligations, a firm may not switch over from one source of labour or row material to the other one.

Similarly the funds placed in bank deposits over a number of years are locked in and even though reasonably appreciable rate of return is available in the money market, due to other modes of investments, persons may not suddenly jump in for such an investment. This is attributed to the problem of security and liquidity of the investments. This can generate Lag model.

6. Adhoc Estimation of DL models

Let us rewrite the finite DL model as

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_K X_{t-K} + U_t$$

A simple approach for estimation is as under

(1) First regress Y_t on X_t, then Y_t on X_t and X_{t-1}, then Y_t on X_t, X_{t-1}, X_{t-2} and So on. (Thus taking 2 variables, then 3 variables and So on)

(2) This procedure is carried out until the last variable turns out to be insignificant and /or the coefficient of at least one of the variables changes sign from positive to negative and vice versa. This procedure appears to be simple but if has many draw backs.

(1) There is no indication beforehand as to the maximum length of the lag.

(2) The model becomes complex to understand and estimate.

(3) The degrees of freedom reduces step by step and hence the conclusions can be unreliable.

(4) In economic time series, the successive lag values are highly correlated. This generates multicollinearity which leads to inefficient estimation. Standard errors increase and this can lead to tests turning out to be insignificant.

7. Koyck's Approach to DL models

Let us consider Infinite DL model as under

$Y_t = \alpha + (\beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots \text{ up to } \infty) + U_t$ .... (1)

Koyck assumed that the subsequent effects of different past periods will reducing subsequently on the current year endogenous variable.

Koyck assumed that all Beats are of the same sign and they are related by

$\beta_K = \beta_0 \lambda^k$ .... (2)

$0 < \lambda < 1,\ k = 0, 1, 2 \ldots$

This assumes that Beta coefficients are declining geometrically

Here λ is called rate of decline or decay of the DL.

$1 - \lambda$ is called Speed of Adjustment.

Here $\beta = \sum_{K=0}^{\infty} \beta_K = \frac{\beta_0}{(1-\lambda)}$ .... (3)

From (1) and (2),

$Y_t = \alpha + (\beta_0 X_t + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \cdots + \beta_0 \lambda^K X_{t-K} + \cdots \text{ up to } \infty) + U_t$

then

$Y_{t-1} = \alpha + (\beta_0 X_{t-1} + \beta_0 \lambda X_{t-2} + \beta_0 \lambda^2 X_{t-3} + \cdots + \beta_0 \lambda^k X_{t-K-1} + \cdots \text{ up to } \infty) + U_{t-1}$

Hence $Y_t - \lambda Y_{t-1} = \alpha(1-\lambda) + \beta_0 X_t + (U_t - \lambda u_{t-1})$ .... (4)

thus $Y_t - \lambda Y_{t-1} = \alpha(1-\lambda) + \beta_0 X_t + \vartheta_t$ ....(5)

Where $\vartheta_t = U_t - \lambda U_{t-1}$ .... (6)

This is called Koyck's transforulation of IDLM into Autoregressive model of the first order.

This transformation becomes much simpler than the original model. Note that here ϑ_t is the moving average of U_t and U_{t-1}.

Here all lagged variables $X_{t-1}, X_{t-2}, \ldots$ etc. are replaced by a single term Y_{t-1},thus multicollinearity is resolved and we have to estimate only α, λ and β_0, which becomes simpler as compared with estimating original IDLM.

8. Some features of Koyck's transformation

(1) Koyck transformation leads to convert an infinite DL model into an autoregressive model.

(2) Since the term Y_{t-1} appears on the right hand side, we have a stochastic explanatory variable in the model. The basic assumption in GLM is that explanatory variable should be distributed non stochastic or if stochastic, it should be distributed independently of the stochastic disturbance term.

(3) The original model contains U_t but transformed model contains $\vartheta_t = U_t - \lambda U_{t-1}$.

Thus if original $U's$ are serially uncorrelated, the $\vartheta's$ are serially correlated.

This generates serial correlation (autocorrelation) problem in addition to the stochastic explanatory variable Y_{t-1} appearing in the right hand side of the final equation.

Note: To deal with this situation we have to develop an alternative to test for serial correlation in the presence of lagged Y. This is given by Durbin h test in place of usual DW test which is discussed below.

9. Detection of Autocorrelation in AR models

DURBIN h TEST

Durbin has proposed a larger sample test of first order serial correlation in AR model.

This is called Durbin's h test

Define $h = \hat{\rho}\sqrt{\frac{n}{1-n[Var(\hat{\alpha}_2)]}}$ .... (1)

Where $n =$ Sample size, $Var(\hat{\alpha}_2) =$ Variance of coefficient α_2 of the term Y_{t-1}.

$\hat{\rho} =$ Estimate of the first order Serial correlation ρ.

An approximate estimate of ρ as given by DW test is $\hat{\rho} \approx 1 - \frac{d}{2}$

Where $d =$ usual DW statistic.

Hence above equation (1) can be written as

$h \doteqdot \left(1 - \frac{d}{2}\right) . \sqrt{\frac{n}{1-n[V(\hat{\alpha}_2)]}}$ .... (2)

Durbin has shown that if $\rho = 0$, h statistic follows standard normal distribution for large n that is $h \sim N(0,1)$

Hence the test can be worked out as under

(1) Consider the regression $Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + +\beta_2 X_{t-2} + \vartheta_t$ by using OLS method and estimate $\alpha_0, \alpha_1, \alpha_2$.

(2) Obtain $V(\hat{\alpha}_2)$

(3) Compute $\hat{\rho}$ as given by $\hat{\rho} \approx 1 - \frac{d}{2}$ (we have to use DW test statistic d) and hence find $\hat{\rho}$

(4) Compute Durbin's h stastics as given by (2) above.

(5) Apply normality test

$H_0: \rho = 0$ (There is no positive auto correlation)

Vs $H_1: \rho \neq 0$

(6) At 5% level of significance, (i) if $h \leq 1.96 \rightarrow$ H_0 may be accepted.

(ii) if $h > 1.96 \rightarrow$ Reject H_0

<u>Note</u>

(1) We need not bother about number of explanatory variables, only the coefficient of Y_{t-1} is important

(2) Test is not applicable if $nV(\hat{\alpha}_2) > 1$ However this usually will not happen in practice

(3) This is a large sample test. Hence it is not justified for small samples.

2. Some applications of Koyck's model

(1) Adaptive Expectation Model

Let us consider the model as

$Y_t = \beta_0 + \beta_1 X_t^* + U_t$............................ (1)

Where Y_t = Demand for money at time t

X_t^* = Anticipated rate of interest at time t

Here X_t^* is not directly observable.

If is proposed to write X_t^* as a convex combination of X_t and X_{t-1}^* thus we write

$X_t^* = \gamma . X_t + (1-\gamma) X_{t-1}^*$ .... (2)

where $0 < \gamma \leq 1$ is called the <u>coefficient of expectation.</u>

X_t = current year interest rate.

If $\gamma = 1, X_t^* = X_t$ (In a way γ and 1-γ are the weights attached to X_t and X_{t-1}^*)

From (1) and (2) we write

$Y_t = \beta_0 + \beta_1 [\gamma X_t (1-\gamma) X_{t-1}^*] + U_t$

$= \beta_0 + \beta_1 \gamma X_t + \beta_1 (1-\gamma) X_{t-1}^* + U_t$ .... (3)

From (1), $Y_{t-1} = \beta_0 + \beta_1 X_{t-1}^* + U_{t-1}$ So that

$(1-\gamma) Y_{t-1} = (1-\gamma)\beta_0 + (1-\gamma)\beta_1 X_{t-1}^* + (1-\gamma) U_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots. (4)$

Hence (3)–(4) gives on simplification the relation

$Y_t = y\beta_0 + \beta_1 y X_t + (1-\gamma) Y_{t-1} + \vartheta_t$..................... (5)

where $\vartheta_t = U_t - (1-\gamma)\, U_{t-1}$.......................... (6)

Equation (5) is similar to Koyck's model.

(2) Partial Adjustment Model

Consider the model as

$Y_t^* = \beta_0 + \beta_1 X_t + U_t$............................ (1)

where X_t= ouput at time t

Y_t^* = anticipated output at time t

Here Y_t^* is not directly observable

Nerlove gave the partial adjustment hypothesis (for stock adjustment) as

$Y_t - Y_{t-1} = \delta(Y_t^* - Y_{t-1})$.................. (2)

where $0 < \delta \leq 1$ is called the <u>coefficient of adjustment</u>

Here $Y_t - Y_{t-1}$ = Actual change in output

$Y_t^* - Y_{t-1}$= Desired change in output

From equation (2), $Y_t = \delta Y_t^* + (1-\delta)\, Y_{t-1}$............. (3)

Hence from (1) and (3), we get

$Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1-\delta)Y_{t-1} + \delta U_t$..................... (4)

which is similar to Koyck's model.

(3) Combination of Adaptive Expectations and Partial Adjustment Models

Consider the model as

$Y_t^* = \beta_0 + \beta_1 X_t^* + U_t$................... (1)

where Y_t^*= Desired stock of capital

and X_t^* = Expected level of output.

Here both Y_t^* and X_t^* are not directly observable.

We put adaptive expectation hypothesis for X_t^* (writing X_t^* $= \gamma \cdot X_t + (1-\gamma)X_{t-1}^*$) and the partial adjustment method for Y_t^* (writing $Y_t - Y_{t-1} = \delta(Y_t^* - Y_{t-1})$)

which gives after simplification the following form

$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + +\alpha_2 X_{t-2} + \vartheta_t$................... (2)

Where $\alpha_0 = \beta_0\delta\gamma$

$\alpha_1 = \beta_1\delta\gamma$

$\alpha_2 = (1-\gamma) + (1-\delta)$

$\alpha_2 = -(1-\delta)(1-\gamma)$

and $\vartheta_t = \delta[U_t - (1-\gamma)U_{t-1}]$

Equation (2) is like Koyck's model.

Here error term follows a moving average process. This model is linear in $\alpha's$ but non-linear in the original parameters.

The basic equation of the model given in (1) is referred to as Friedman's income hypothesis which states that permanent or long run consumption is a function of permanent or long run income.

11. Estimation of Autoregressive Models

We summarise the above applications (1) and (2) by writing the models

(I) **Koyck** $Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + (U_t - \lambda U_{t-1}) \quad \rightarrow (1)$

(II) **AEM** $Y_t = \gamma\beta_0 + \gamma\beta_1 X_t + (1-\gamma)Y_{t-1} + (U_t - (1-\gamma)U_{t-1}) \quad \rightarrow (2)$

(III) **PAM** $Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1-\delta)Y_{t-1} + \delta U_t \quad \rightarrow (3)$

All the above three representations of the model can be stated by means of a common equation

$$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + \vartheta_t \quad \rightarrow (4)$$

Where $\alpha_0, \alpha_1, \alpha_2$ and ϑ_t have the meanings and the relations with the respective models if we compare them individually with equation (4).

All the above system models are autoregressive in nature. Ordinary least squares method is not applicable due to stochastic nature of the term Y_{t-1} on the right hand side and also due to the possibility of serial correlation due to the term ϑ_t.

However if it can be shown that Y_{t-1} is independently distributed of ϑ_t then OLS method can be applied.

For Koyck's model, $\vartheta_t = U_t - \lambda U_{t-1}$

$E(\vartheta_t) = 0$ for all t

and $cov(\vartheta_t, \vartheta_{t-1}) = -\lambda\sigma^2 \neq 0$

Also $cov(Y_{t-1}, \vartheta_t) = -\lambda\sigma^2 \neq 0$

Similar results hold good for AEM and PAM also. Thus if we apply OLS method, the results will be misleading. In the case of PAM, $\vartheta_t = \delta U_t$ .Thus basic assumption of classical linear model is satisfied. Here OLS estimators will be consistent but they are biased.

Due to all these reasons, OLS method should not be used directly for estimation of AR models. Instead of this, some simpler methods should be devised.

12. Method of Instrumental Variables to estimate Koyck's model

We consider the general equation (Koyck's model) as

$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + \vartheta_t$.................. (1)

This is AR form obtained on transformation of DL model.

If we apply straight way OLS method of equation (1)

We get the following normal equations

$\sum Y_t = n\hat{\alpha}_0 + \hat{\alpha}_1 \sum X_t + \hat{\alpha}_2 \sum Y_{t-1}$ .... (2)

$\sum X_t Y_t = \hat{\alpha}_0 \sum X_t + \hat{\alpha}_1 \sum X_t^2 + \hat{\alpha}_2 \sum X_t Y_{t-1}$ .... (3)

$\sum Y_{t-1} Y_t = \hat{\alpha}_0 \sum Y_{t-1} + \hat{\alpha}_1 \sum X_t Y_{t-1} + \hat{\alpha}_2 \sum Y_{t-1}^2$ .... (4)

The solutions obtained from equations (2), (3) and (4) may not give consistent estimators due to Y_{t-1} and ϑ_t which are correlated, although X_t and X_{t-1} may not be correlated with ϑ_t.

To resolve this problem Liviatan used Instrumental variables Method (IVM). He used X_{t-1} as a proxy Variable in place of Y_{t-1}. Thus equation (2) & (3) remain unchanged but equation (4) changes and the new set of normal equations are written as under

$\sum Y_t = n\hat{\alpha}_0 + \hat{\alpha}_1 \sum X_t + \hat{\alpha}_2 \sum Y_{t-1}$ .... (5)

$\sum Y_t X_t = n \sum X_t + \hat{\alpha}_1 \sum X_t^2 + \hat{\alpha}_2 \sum Y_{t-1} X_t$ .... (6)

$\sum Y_t X_{t-1} = \hat{\alpha}_0 \sum X_{t-1} + \hat{\alpha}_1 \sum X_t X_{t-1} + \hat{\alpha}_2 \sum Y_{t-1} X_{t-1}$ .... (7)

(Only equation 7 changes from equation (4))

Livitan has shown that solving (5), (6) and (7) gets consistent estimators, which is done by means od the proxy variable X_{t-1} in place of Y_{t-1} (in the 3rd equation).

These estimators are called **<u>IV estimators</u>**. Here there can be problem of multicollinearity and hence the estimators are likely to be inefficient.

3. Some preliminary concepts about Lag structure

Here we want to introduce a lag operator which can handle lag structure in the case of polynomial lag and there by define Mean lag, Median lag, triangular lag etc.

Let us consider DL model as given by

$Y_t = \mu + \delta_0 X_t + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \delta_3 X_{t-3} + U_t$......... (1)

In this model, one does not know in advance about the maximum length of lag.

We can write a general relation as

$Y_t = \mu + D(L) \cdot X_t + U_t$.......................... (2)

Where D(L) is a polynomial of some degree S in the lag operator.

That is $D(L) = \delta_0 + \delta_1 L + \cdots + \delta_s L^s$....................... (3)

In practice, the maximum lag may have to be fairly large to provide an adequate presentation of the relationship between Y in terms of X's.

We define Mean Lag and Median Lag as under

<u>Mean Lag</u>: It is the length of time it takes for a unit change in the explanatory variable to the dependent variable Y.

<u>Median Lag</u> : Median Lag is the time required for the first half of the total change in Y following a unit sustained change in X.

$\left(\text{Based upon above relation (3), it can be shown that Mean Lag } = \frac{D'(1)}{D(1)}\right)$....

(4)

Ex. Show that for IDL model with Koyck transformation,

$\textbf{Mean Lag} = \frac{\lambda}{(1-\lambda)}$

We consider IDL model as

$Y_t = \mu + \delta_0 X_t + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \ldots + U_t$.................... (1)

Kyock assumes $\delta_k = \delta_0 \lambda^k$, $k = 0, 1, 2$ ...., $0 < \lambda < 1$

Thus

$Y_t = \mu + \delta_0 X_t + \lambda\delta_0 X_{t-1} + \lambda^2 \delta_0 X_{t-2} + \ldots + U_t$.................. (2)

Here $D(L) = \delta_0 + \lambda\delta_0 + \lambda^2\delta_0 +$........................ (3)

So that $Y_t = \mu + D(L)X_t + U_t$.......................... (4)

$\text{Mean Lag} = \frac{D'(1)}{D(1)}$

Here $D(L) = \delta_0 + \lambda\delta_0 + \lambda^2\delta_0 + \cdots$

$\therefore D'(L) = 1 \cdot \delta_0 + 2 \cdot \lambda\delta_0 + \cdots$

$\therefore D'(1) = 1 + 2\lambda + 3\lambda^2 + \cdots$

$= \frac{\lambda}{(1-\lambda)^2}$

$D(1) = 1 + 2\lambda + 3\lambda^2 + \cdots = \frac{1}{(1-\lambda)}$

$\therefore \text{ Mean Lag} = \frac{D'(1)}{D(1)} = \frac{\lambda}{(1-\lambda)}$

$\left[Note\ It\ can\ be\ shown\ that\ for\ Koyck's\ model, Meadian\ Lag = -\left(\frac{\log 2}{\log \lambda}\right)\right]$

Triangular or Arithmetic Lag Structure

Here the model assumes that the explanatory variable exerts its greatest impact in the current period and then it declines by equal increments to zero as one goes in the distance past.

We can consider several regressions as under representing triangular or arithmetic lag structure

e.g. $Y_t = \alpha + \beta\left(\frac{2X_t + X_{t-1}}{3}\right) + U_t$

$$Y_t = \alpha + \beta\left(\frac{3X_t + 2X_{t-1} + X_{t-2}}{6}\right) + U_t$$

$$Y_t = \alpha + \beta\left(\frac{4X_t + 3X_{t-1} + 2X_{t-2} + X_{t-3}}{10}\right) + U_t$$ and so on.

We can run such several regressions and them choose the one as the best for which R^2 is highest.

Note

1) We can have extended studies based upon Almon's polynomial lag approach etc.
2) In our discussion in this lecture, we have omitted rigorous mathematical treatment and other related test statistics looking to the limitations of our syllabus.

But there is a large category of studies related to Lag models which are very useful and interesting.