ESTIMATING PARAMETERS IN AUTOREGRESSIVE PROCESS DURING ROLLING PERIODS.

Phongsaphen Jantana*
Prapaisri Sudasna-na-Ayudthya

Department of Industrial Engineering
Faculty of Engineering, Kasetsart University, Bangkok, Thailand
Graduate student, Industrial Engineering Department, Kasetsart University
Assistant Professor in Industrial Engineering Department at Kasetsart University

ABSTRACT

There are many techniques used to estimate autoregressive (AR) model parameters, but they have disadvantages such as high calculation times and errors. This study compared operation times and prediction mean square error during rolling periods between two different methods. One was the least squares (LS) method, and the other was the Yule-Walker (YW) method. After the linear system equations were obtained by the least squares method, they were solved and updated using invertibility method. The other linear system equations were obtained by the Yule-Walker method and they were solved using Durbin-Levinson Recursion and invertibility method. The study showed that the first method is the preferred technique for long rolling periods because it has less calculation time and it has same prediction mean square error.

KEYWORDS: Autoregressive process, Estimating parameters, Least squares, Yule-Walker.

1. INTRODUCTION

There are many variables important for the decision process in industry, such as the demand for managing production and inventory systems. The ability to analyze historical data to predict future events should reduce the risk in decision-making. Many industrial data are highly dependent on observations. Therefore, the autoregressive process was developed to obtain the relationship of time-series data within itself [1]. One of the pioneering works is the AR (2) model for sunspot data that analyzed by G.U.Yule in 1927. In 1931 Gilbert Walker considered the case AR (4) model for Dawin Pressure [8]. Box and Jenkins developed a well-known methodology for AR model in 1970[5]. They used an estimating method equivalent to the least squares method [6]. In 1993, Janacek and Swift studied about the estimating method of AR(p) model parameters that have assumed a zero mean Model [4]. They found that for the parameters were obtained for one time of one-step-ahead prediction, the Yule-Walker method that used Durbin-Levinson Recursion was quicker than the least squares estimates [2].

]In this study, the parameters were obtained from a univariate stationary autoregressive process. The study considered population mean of the process in three different situations and the autoregressive process used three methods to estimate parameters. First method was least squares method with invertibility, second was the Yule-Walker method with Durbin-Levinson Recursion, and the last one was Yule-Walker method with invertibility.

When periods roll for a long time, the parameters of a time series model must be changed for updating the model. Which method should be preferred to obtaine the model parameters ?

The purpose of this study was to implement an approach for using the invertibility method for forecasting the autoregressive process. The invertibility method is a numerical approach that can solve linear systems equations from the least squares method and the Yule-Walker method. When the linear systems equation shifts, invertibility method can update these systems equations [2]. Cost and accuracy of forecasting are two aspects to consider for the implementation of any forecasting method.

* Corresponding author :
E-mail: g4485009@ku.ac.th and fengpsa@ku.ac.th

2. METHODOLOGY

In this study, linear stochastic stationary time series data were fitted to an autoregressive model by two methods. One was the least squares method, and the other was the Yule-Walker method. The Least Squares method fitted a model by minimizing the sum of square errors for estimating parameters. It used the normal equations to implement the linear system. This method solved the linear system by invertibility method. The Yule-Walker method used autocorrelation to formulate the Yule-Walker equation, which solved model parameters by Durbin-Levinson Recursion and invertibility methods.

This study considered the population mean of the process in three different situations. First, the process does not obtain the population mean. Second, the process's mean is zero, and third, the process's mean is not zero. The two methods can be described as:

2.1 The least squares method

The AR(p) model of order $p \geq 1$ for all $t > p$ when the process's mean is not obtained is shown in equation (1).

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \varepsilon_t \qquad (1)$$

Where x_t is an observation at time t and $\hat{\phi}_0$ is the scalar while $\hat{\phi}_i$ are the unknown autoregressive parameters. When $\hat{x}_t$ is a forecasting value at time t, have an error ε_t is distributed as N(0 , σ^2). When minimizing the sum of square errors for estimating parameters, it gave the linear systems equation $Ax = b$ from least squares normal equation as follows:

$$\begin{bmatrix} \sum_{t=p+1}^{T} 1 & \sum_{t=p+1}^{T} x_{t-1} & \cdots & \sum_{t=p+1}^{T} x_{t-p} \\ \sum_{t=p+1}^{T} x_{t-1} & \sum_{t=p+1}^{T} x_{t-1}^2 & \cdots & \sum_{t=p+1}^{T} x_{t-1}x_{t-p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=p+1}^{T} x_{t-p} & \sum_{t=p+1}^{T} x_{t-1}x_{t-p} & \cdots & \sum_{t=p+1}^{T} x_{t-p}^2 \end{bmatrix} \begin{bmatrix} \hat{\phi}_0 \\ \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} = \begin{bmatrix} \sum_{t=p+1}^{T} x_t \\ \sum_{t=p+1}^{T} x_t x_{t-1} \\ \vdots \\ \sum_{t=p+1}^{T} x_t x_{t-p} \end{bmatrix} \qquad (2)$$

When "***A***" is a square matrix formed by using sum of backward elements at a time and "A^{-1}" is the inverse of "***A***" while "***x***" is a parameter vector and "***b***" is a right hand side vector, and "***T***" is the end of observation time, the linear systems were solved using the invertibility method [2]. It is define as:

$$A_T x_T = b_T$$
$$A_T^{-1} A_T x_T = A_T^{-1} b_T \qquad (3)$$

The results from this method were used as the model parameters. The AR(p) models were then ready to be used to forecast for time series. After forecasting, the forecast errors were computed to compare the prediction mean square error, to consider forecast accuracy.

Time series data were always shifted to the next period because the parameters must have new estimated values to update the model. The inverse matrix A_T^{-1} was updated by using Sherman-Morison-Wolbury method [2]. The formula is shown in equation (4).

$$A_{T+1}^{-1} = (A_T + uv')^{-1} = A_T^{-1} - \frac{A_T^{-1} u v' A_T^{-1}}{1 - v' A_T^{-1} u} \qquad (4)$$

Where $A_{T+1} = A_T + uv'$ and "***u***" and "***v***" are vectors. This method did not need to know the population mean of the process. However, another situation did need to know the population mean of the process. When $y_t = x_t - \mu$ then the AR(p) model is defined as follows:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \tag{5}$$

In this study, the time series data were fitted to an autoregressive model in equation (5) when the process mean was known. Then, the second situation is $\mu = 0$ and the last one is $\mu \neq 0$ and, the equation in (2) did not have the first row and the first column of matrix A. However, the time series data were able to fit to an autoregressive model in equation (5) by the Yule-Walker method.

2.2 The Yule-Walker method

The Yule-Walker method gives the linear systems equation $Ax = b$ as follows [3]:

$$\begin{bmatrix} 1 & r_1 & \cdots & r_{p-1} \\ r_1 & 1 & \cdots & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix} \tag{6}$$

The population correlation coefficient (ρ_k) was estimated by the sample correlation coefficient (r_k) while $k > 0$ is the time lag. Then "***A***" is a square matrix formed by using sample correlation coefficient (r_k), while "***x***" is a parameter vector and "***b***" is a right hand side vector of correlation coefficient.

The linear systems were solved using Durbin-Levinson Recursion that defined as follows [4]:

$$\hat{\phi}_{kk} = \frac{r_k - r_{k-1}\hat{\phi}_{1,k-1} - r_{k-2}\hat{\phi}_{2,k-1} - \ldots - r_1\hat{\phi}_{k-1,k-1}}{1 - r_{k-1}\hat{\phi}_{k-1,k-1} - r_{k-2}\hat{\phi}_{k-2,k-1} - \ldots - r_1\hat{\phi}_{1,k-1}} \tag{7}$$

When time series data were shifted to the next period, in Durbin-Levinson Recursion, the new parameters were obtained by repeatedly calculating the new correlation coefficient (r) and the parameters in equation (7). In the other way, the linear system in equation (6) was could be solved using the invertibility method as defined in equation (3). The parameters were updated by the Sherman-Morison-Wolbury formula as in equation (4).

In this study, the time series data were fitted to an autoregressive model in equation (5). When the process did not know the population mean, then the sample mean ($\bar{x}$) was used to estimate the population mean (μ). When the process mean was known, the second situation was $\mu = 0$ and the third was $\mu \neq 0$.

This study used the generated undamped sinusoids autoregressive data in three orders, namely order 5, order 12 and order 20. First, forecast for one-step-ahead only one time. The study varied the number of history data from 5,000, 10,000, 15,000, 20,000, 25,000 to 30,000 observations. Second, forecast for one-step-ahead when periods were rolled. Forecast used the first 15,000 observations, then periods were rolled 1,000, 5,000, 10,000, 15,000, 20,000 to 25,000 times. Every method in this study was evaluated for operation time as cost of computing and used prediction mean square error for accuracy of the forecast methods. MATLAB was used to generate data and compute all programmes.

3. RESULTS AND DISCUSSION

This study showed the results for one-step-ahead of one time forecasting and one-step-ahead of long rolling periods forecasting in each method. One was the least squares and the other was the Yule-Walker, to compared operation time and prediction error when the population means were obtained and were not obtained. The results for one-step-ahead of one time forecasting are shown in Tables 1and 2. The study showed the forecast for AR (5) that varied the number of history data. The study found that for one-step-ahead of one time forecasting, the Yule-Walker equation with Durbin-Levinson Recursion had higher efficiency than the least squares with invertibility.

Next, the results for one-step-ahead of long rolling periods forecasting are shown in Tables 3 and 4. The tables show the results of AR(5), AR(12)and AR(20), with forecast started at 15,000th observation and forecast one-step-ahead. Then the periods were rolled for 1,000 times, 5,000 times, 10,000 times, 15,000 times, 20,000 times and 25,000 times, the autoregressive process was forecasted one-step-ahead for all those periods.

After that, Figures1,2 and 3 compares the operation time of two forecasting methods. One is least squares with update invertibility when the population mean was not obtained. The other is the Yule-Walker equation with Durbin-Levinson Recursion when $\mu = 0$. The study showed that the results for a long time of rolling periods of the least squares method with the update invertibility method was obtained faster than other methods and there were the same prediction mean squares errors from forecasting (see table 4). The least squares method can update directly when the period was rolled. It did not need to repeat the calculated parameters. While Yule-Walker used Durbin-Levinson Recursion and the invertibility method, the Durbin-Levinson Recursion obtained the parameters faster than the invertibility method, because the invertibility method consumed operation time when repeating computation of the sample correlation coefficient (r_k).

Table 1 Comparison of time of calculations to estimate AR(5) parameters obtained for one time when using different methods.

AR (p)	Method \ No.of periods			Time(sec.) 5000	10000	15000	20000	25000	30000
AR (5)	Least Squares	inverse	ϕ_0	0.420	0.551	0.681	1.152	2.203	4.446
			$\mu=0$	0.441	0.520	0.681	1.122	2.483	4.566
			$\mu\neq 0$	0.511	0.721	1.041	1.682	3.945	8.352
	Yule-Walker	inverse	$\mu=\bar{x}$	0.260	0.341	0.500	0.731	2.123	4.176
			$\mu=0$	0.160	0.170	0.181	0.190	0.211	0.230
			$\mu\neq 0$	0.261	0.341	0.501	1.131	2.243	4.306
		Durbin-Levinson	$\mu=\bar{x}$	0.140	0.231	0.391	1.302	2.153	4.717
			$\mu=0$	0.051	0.060	0.070	0.080	0.091	0.111
			$\mu\neq 0$	0.120	0.220	0.380	0.871	1.963	4.136

Table 2 Comparison of prediction square errors obtained from fitting AR(5) model to the data using different methods.

AR (p)	Method \ No. of periods			Prediction Square Error 5000	10000	15000	20000	25000	30000
AR (5)	Least Squares	inverse	ϕ_0	0.0021	0.0247	0.000158	0.0216	0.0034	0.0528
			$\mu=0$	0.0020	0.0247	0.000179	0.0215	0.0035	0.0526
			$\mu\neq 0$	0.0020	0.0247	0.000179	0.0215	0.0035	0.0526
	Yule-Walker	inverse	$\mu=\bar{x}$	0.0021	0.0246	0.000156	0.0216	0.0034	0.0528
			$\mu=0$	0.0020	0.0247	0.000179	0.0215	0.0035	0.0526
			$\mu\neq 0$	0.0020	0.0247	0.000179	0.0215	0.0035	0.0526
		Durbin-Levinson	$\mu=\bar{x}$	0.0021	0.0246	0.000156	0.0216	0.0034	0.0528
			$\mu=0$	0.0020	0.0247	0.000179	0.0215	0.0035	0.0526
			$\mu\neq 0$	0.0020	0.0247	0.000179	0.0215	0.0035	0.0526

Table 3 Comparison of time of calculations to estimate parameters obtained for long rolling periods when using different methods.

AR (p)	Method		No.of rolling periods	Time(sec.) 1000	5000	10000	15000	20000	25000
AR (5)	Least Squares	inverse	ϕ_0	0.791	1.061	1.232	1.463	1.702	1.913
			$\mu=0$	1.032	1.322	1.572	1.802	1.952	2.193
			$\mu\neq0$	1.272	2.333	4.547	8.052	11.637	16.203
	Yule-Walker	inverse	$\mu=\bar{x}$	7.951	42.542	96.970	163.335	241.357	332.969
			$\mu=0$	1.222	4.367	10.265	17.074	24.835	34.770
			$\mu\neq0$	0.661	2.053	4.817	8.703	12.649	17.345
		Durbin-Levinson	$\mu=\bar{x}$	6.529	35.271	80.576	135.886	200.118	277.088
			$\mu=0$	0.180	0.901	2.594	5.668	8.522	12.088
			$\mu\neq0$	0.831	3.235	4.777	15.191	22.493	31.636
AR (12)	Least Squares	inverse	ϕ_0	0.841	1.192	1.482	1.643	1.903	2.283
			$\mu=0$	1.232	1.492	1.622	1.903	2.183	2.574
			$\mu\neq0$	1.262	2.543	4.717	8.302	11.556	16.414
	Yule-Walker	inverse	$\mu=\bar{x}$	14.460	79.625	182.142	304.678	450.648	620.362
			$\mu=0$	0.882	3.145	6.900	11.907	16.875	22.933
			$\mu\neq0$	1.462	5.689	12.388	21.331	30.494	40.879
		Durbin-Levinson	$\mu=\bar{x}$	13.450	72.264	164.497	277.238	410.630	566.315
			$\mu=0$	0.210	0.801	2.924	6.129	9.323	13.069
			$\mu\neq0$	0.841	2.093	5.658	15.462	23.263	31.155
AR (20)	Least Squares	inverse	ϕ_0	0.962	1.202	1.573	2.303	2.493	2.594
			$\mu=0$	1.372	1.602	1.952	2.193	2.553	2.875
			$\mu\neq0$	1.412	2.644	4.676	8.893	12.688	16.513
	Yule-Walker	inverse	$\mu=\bar{x}$	22.342	122.436	278.301	468.914	692.927	950.857
			$\mu=0$	1.242	5.167	10.936	18.086	24.805	33.606
			$\mu\neq0$	1.973	7.941	16.654	27.329	38.795	50.683
		Durbin-Levinson	$\mu=\bar{x}$	20.510	113.744	259.914	439.452	649.945	895.768
			$\mu=0$	0.280	1.121	3.625	7.051	10.796	15.011
			$\mu\neq0$	0.731	3.946	9.894	17.135	24.836	34.249

From Table 3, some results were plotted as shown in Figures 1,2 and 3.

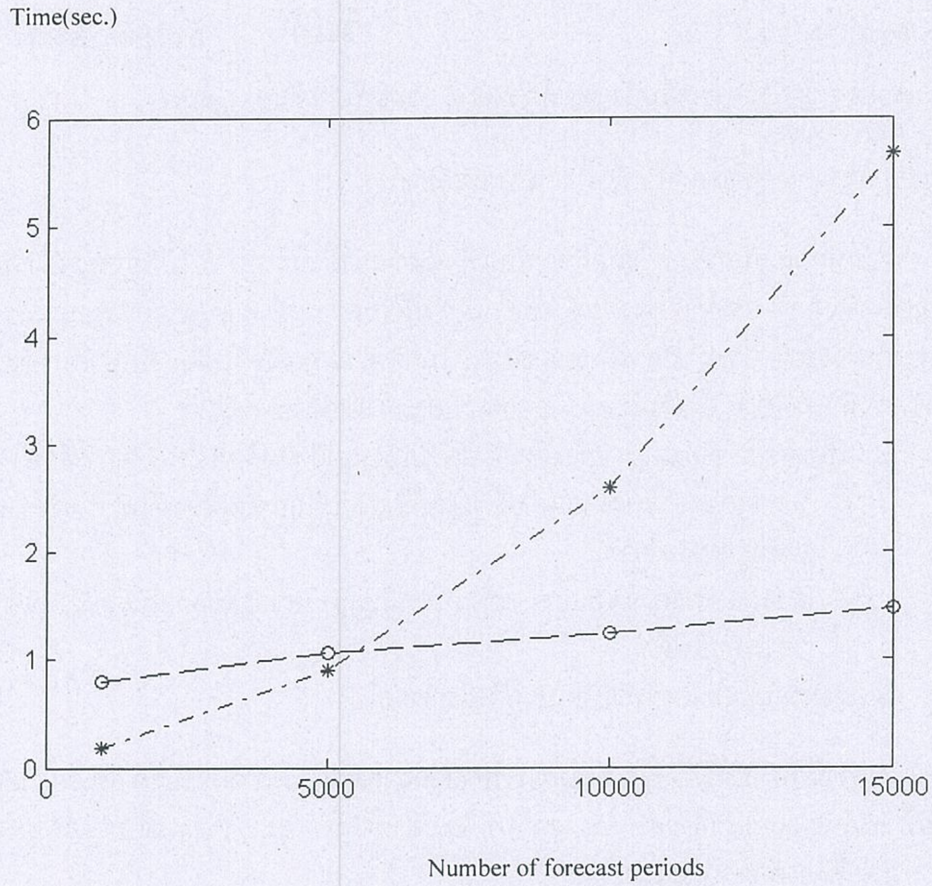

Figure 1 The AR (5) data of time versus number of forecast periods of the least squares method when the process mean is not obtained (- o -) and the Yule-Walker method when the process mean is zero (- * -).

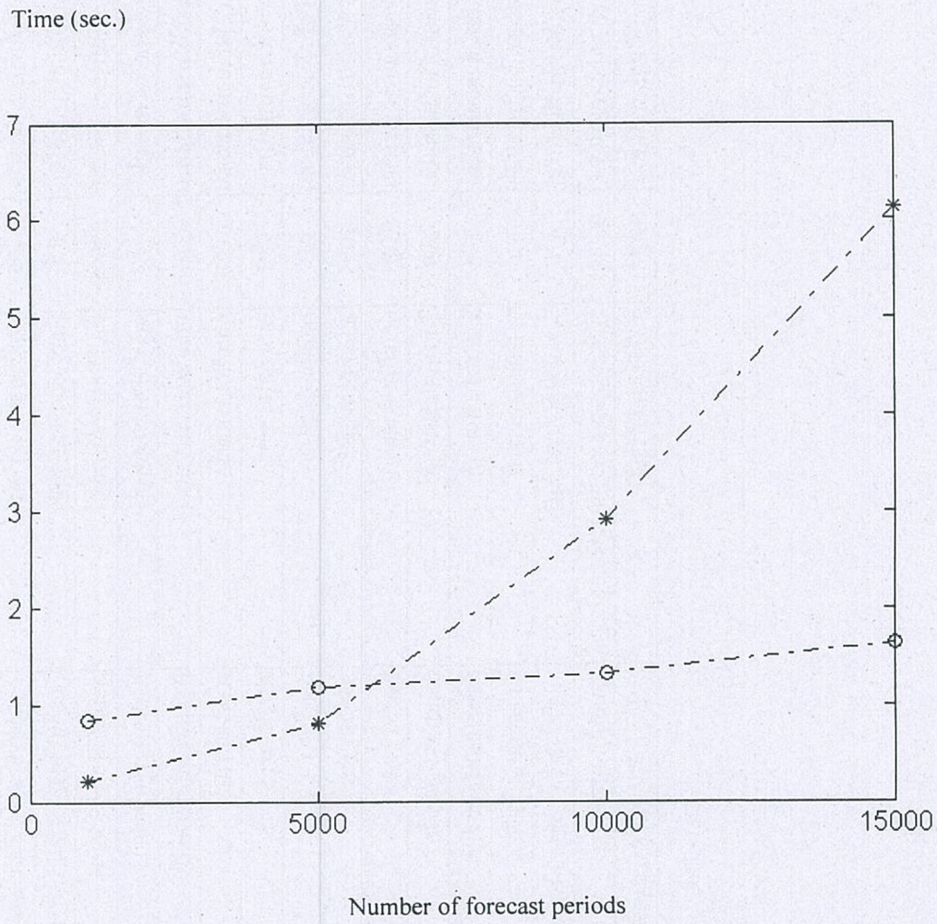

Figure 2 The AR (12) data of time versus number of forecast periods of the least squares method when the process mean is not obtained (- o -) and the Yule-Walker method when the process mean is zero (- * -).

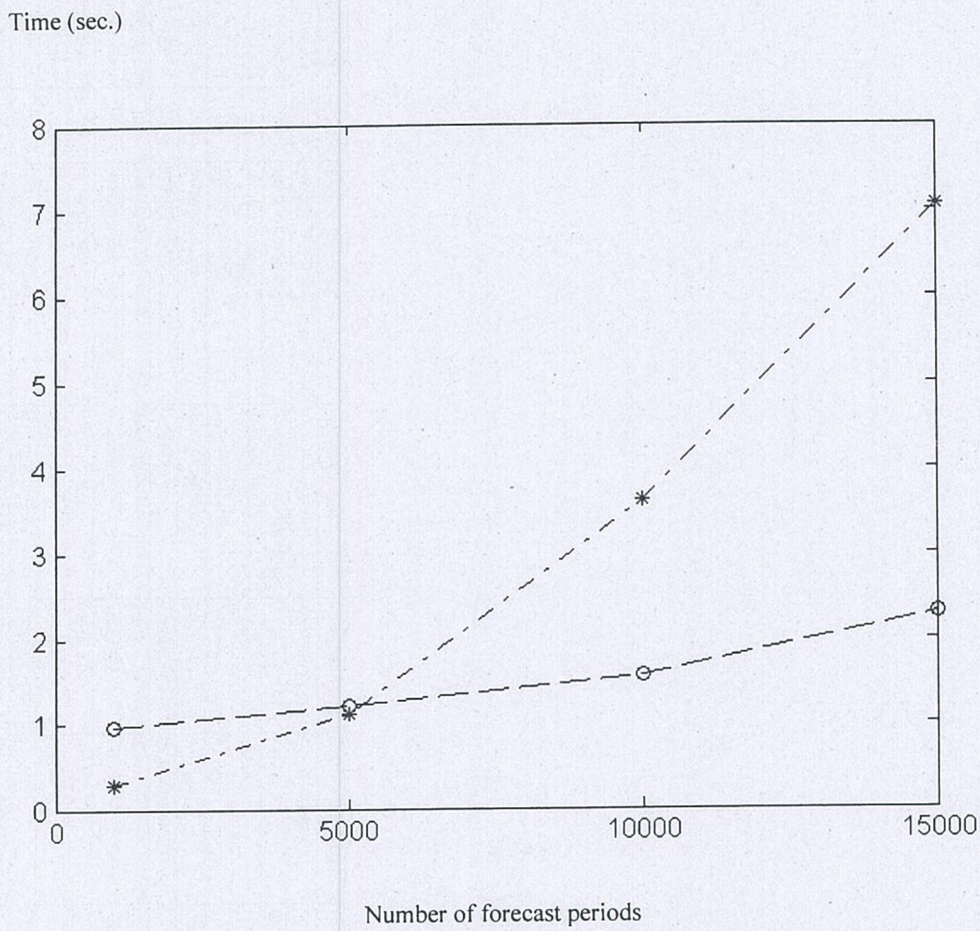

Figure 3 The AR (20) data of time versus number of forecast periods of the least squares method when the process mean is not obtained (- o -) and the Yule-Walker method when the process mean is zero (- * -).

Table 4 Comparison of prediction means square errors obtained from fitting the AR (p) model to the data using different methods.

AR (p)	Method \ No. of rolling periods			Prediction Means Square Error 1000	5000	10000	15000	20000	25000
AR (5)	Least Square	inverse	ϕ_0	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
			$\mu=0$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
			$\mu\neq0$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
	Yule-Walker	inverse	$\mu=\overline{x}$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
			$\mu=0$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
			$\mu\neq0$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
		Durbin-Levinson	$\mu=\overline{x}$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
			$\mu=0$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
			$\mu\neq0$	0.0158	0.0147	0.0146	0.0146	0.0146	0.0146
AR (12)	Least Square	inverse	ϕ_0	0.0109	0.0102	0.0101	0.0101	0.0101	0.0101
			$\mu=0$	0.0109	0.0102	0.0101	0.0101	0.0101	0.0101
			$\mu\neq0$	0.0109	0.0102	0.0101	0.0101	0.0101	0.0101
	Yule-Walker	inverse	$\mu=\overline{x}$	0.0108	0.0102	0.0101	0.0101	0.0101	0.0101
			$\mu=0$	0.0108	0.0102	0.0101	0.0101	0.0101	0.0101
			$\mu\neq0$	0.0109	0.0102	0.0101	0.0101	0.0101	0.0101
		Durbin-Levinson	$\mu=\overline{x}$	0.0108	0.0102	0.0101	0.0101	0.0101	0.0101
			$\mu=0$	0.0109	0.0102	0.0101	0.0101	0.0101	0.0101
			$\mu\neq0$	0.0109	0.0102	0.0101	0.0101	0.0101	0.0101
AR (20)	Least Square	inverse	ϕ_0	0.000687	0.000640	0.000633	0.000634	0.000633	0.000634
			$\mu=0$	0.000687	0.000640	0.000633	0.000634	0.000633	0.000634
			$\mu\neq0$	0.000687	0.000640	0.000633	0.000634	0.000633	0.000634
	Yule-Walker	inverse	$\mu=\overline{x}$	0.000688	0.000640	0.000634	0.000634	0.000634	0.000634
			$\mu=0$	0.000688	0.000641	0.000635	0.000636	0.000635	0.000636
			$\mu\neq0$	0.000688	0.000640	0.000634	0.000634	0.000634	0.000634
		Durbin-Levinson	$\mu=\overline{x}$	0.000688	0.000641	0.000634	0.000634	0.000634	0.000634
			$\mu=0$	0.000688	0.000640	0.000634	0.000634	0.000634	0.000634
			$\mu\neq0$	0.000688	0.000640	0.000634	0.000634	0.000634	0.000634

4. CONCLUSION

When forecasting stationary autoregressive process for one-step-ahead of one time forecasting, the Yule-Walker equation with Durbin-Levinson Recursion is the most appropriate method, which always had higher efficiency than the least squares with invertibility. But, when AR (p) model was used to forecast for long rolling periods of stationary process, the least squares with the update invertibility method could estimate the autoregressive model parameters more efficiently. The advantage is that the least squares method with the update invertibility method did not need to obtain the population mean (μ). It is recommended that further study be performed to determine if the least squares technique with the invertibility method can obtain the autoregressive model parameters within one calculation, where the parameters are valid forever, like the direct smoothing in regression [5]. This could be a convenient way to estimate the autoregressive model parameters.

REFERENCES

[1] Box, G.P., Jenkins, G.M., and Reisel, G.C. **1994** *Time Series Analysis Forecasting and Control*. New Jersey, Prentice Hall.

[2] Golub, G.H., and Van Lone, C.F. **1996** *Matrix Computations*. London, The Johns Hopkins Press.

[3] Hannan E.J., Krishaiah, P.R. and Rao, M.M. **1995** *Time Series in Time Domain*, Elsevier Science B.V., Netherlands.

[4] Janacek, G. and Swift, L. **1993** *Time Series Forecasting, Simulation, Applications*. West Sussex, Ellis Horwood Limited.

[5] Montgomery, D.C., Johnson, L.A. and Gardiner, J.S. **1990** *Forecasting and Time Series Analysis*. Singapore, McGraw-Hill, Inc.

[6] Zioutas, G.,Camariopoulos, L. and Bora Senta, E. **1997** *Robust autoregressive estimates using quadratic programming*. European Journal of Operational Research 101 (1977). Elservier Science.

[7] www.http://www.google.co.th/search?q=autoregressive

[8] www.http://www.stat.wharton.upenn.edu/~steele/Carses/956/ResourceDet ails/YuleWalker And more.html