A New ariant of Correlation Approach for ARMA Model Identification

Sang Man Seong*

* School of Mechatronics, Korea University of Technology and Education, Cheonan, Chungnam, Korea
(Tel : +82-41-560-1313; E-mail: ssman@kut.ac.kr)

Abstract: We proposed a new variant of correlation approach for ARMA model. The proposed method is is intended to make the current prediction error uncorrelated with the past one. In the investigation of the properties, the uniqueness, consistency and asymptotic normality of the estimate are shown. Via simulation results, we show that the proposed method give good estimates for various systems.

Keywords: correlation approach, ARMA model, identification, PLR method

1. INTRODUCTION

The identification methods based on correlation approach such as IV(Instrumental Variable) methods and PLR(Pseudo Linear Regression) methods have been studied intensively for several decades. The IV method originally can be applied to estimate the AR(Auto Regressive) part when we consider the ARMA model. On the IV methods, many variants and the properties such as consistency and accuracy were studied in [1]. To improve the parameter estimation accuracy, the over-determined recursive IV methods were suggested in [2-3]. More recently, a method for the identification of closed loop system was proposed in [4].

As another correlation approach, the PLR methods can be applied to estimate both the AR and MA(Moving Average) parts. The accuracy properties including asymptotic normality of a general PLR are studied in [5]. For the general PLR method and modified one, the local convergence analysis was done in [6].

In this paper, we propose a new variant of correlation approach for the estimation both the AR and MA parts of ARMA model. The key idea of this method is to find the estimate making the prediction error at current time become uncorrelated with the one at past time. This is because we call the proposed method as DWM(Direct Whitening Method). As the properties which are required in good parameter estimation methods, we prove the uniqueness, consistency, and asymptotic normality of the estimate acquired by DWM. Even if DWM is an extension of PLR, DWM requires nonlinear optimization differently from PLR. With some simulation results, the performance of DWM is verified for various systems.

2. PRELIMINAR

In this section, model structures and the general correlation approach are explained. Consider ARMA(p, q) model.

$$\ _t + a_1 \ _{t-1} + \cdots + a\ \ _{t-} = e_t + c_1 e_{t-1} + \cdots + c\ e_{t-}$$

where a_i's and c_i's are the parameters, $\ _t$ is the output, and e_t is the white Gaussian noise with zero mean and variance of σ^2. For later use, let's define

$$A(z) := 1 + a_1 z^{-1} + \cdots + a\ z^{-}$$

$$(z) := 1 + c_1 z^{-1} + \cdots + c\ z^{-}$$

$$\theta^* := \begin{bmatrix} a_1 & \cdots & a \end{bmatrix}^T \text{ or}$$

$$\theta^* := \begin{bmatrix} a_1 & \cdots & a & c_1 & \cdots & c \end{bmatrix}^T$$

where $A(z)$ is AR part, (z) is MA part and θ^* is true parameter vector.

The different definition of θ^* depends on what parameters are estimated by a method. We will use the notation $\hat{\theta}$ for the estimate of θ^* and θ for the candidate of $\hat{\theta}$. Let's assume that $A(z)$ and (z) have zeros inside the unit circle respectively and, moreover, both has no common factors. The prediction error is defined as $\varepsilon_t = \ _t - \hat{\ }_{t/t-1,\theta}$. $\hat{\ }_{t/t-1,\theta}$ is the estimate of $\ _t$ at the time $t-1$. Since ε_t is a function of θ, $\varepsilon(t,\theta)$ is used alternatively in this paper.

The correlation approach is based on the idea that the prediction error should be independent of the past output. Therefore the parameter estimate by correlation approach is given by equation (1)[7].

$\hat{\theta} =$ the solution the equation

$$\left[\frac{1}{N} \sum_{t=1}^{N} \zeta(t,\theta) \alpha(\varepsilon(t,\theta)) = 0 \right] \tag{1}$$

where N is the number of measurements, $\zeta(t,\theta)$ is a function of past output and $\alpha(\varepsilon(t,\theta))$ is filtered prediction error.

In standard PLR method, the $\zeta(t,\theta)$ is

$$\zeta(t,\theta) = \begin{bmatrix} \ _{t-1} & \cdots & \ _{t-} & \varepsilon_{t-1} & \cdots & \varepsilon_{t-} \end{bmatrix}^T$$

and $\alpha(\varepsilon(t,\theta)) = \varepsilon(t,\theta)$. To solve the equation (1), least squares method can be used. Then the estimate of standard PLR is given by

$$\hat{\theta} = \left[\frac{1}{N} \sum_{t=1}^{N} \zeta(t,\theta) \zeta^T(t,\theta) \right]^{-1} \frac{1}{N} \sum_{t=1}^{N} \zeta(t,\theta)\ \ _t \tag{2}$$

Since the $\hat{\theta}$ may not satisfy that $A(z)$ and (z) have zeros inside the unit circle. the projection into the unit circle is required. The projection algorithm can be a good method for this purpose[9].

3. DWM AND ITS PROPERTIES

In this section, DWM is presented and its properties are derived. At first, we will show that if we acquire the estimates making the prediction error at current time become uncorrelated with the one at past time, the estimate become the true parameter as the sample number goes to infinite.

Now let's $\zeta(t,\theta)$ in equation (1) as

$$\zeta(t,\theta) = \left[\varepsilon_{t-1} \quad \cdots \quad \varepsilon_{t-} \quad \varepsilon_{t- -1} \quad \cdots \quad \varepsilon_{t- -}\right]^T$$

And $\alpha(\varepsilon(t,\theta)) = \varepsilon(t,\theta)$ like as in standard PLR. Then the estimate is given by

$$\hat{\theta} = \text{the solution of } \left[R_N(k) = 0,\ k = 1, \cdots, \quad + \quad\right] \quad (3)$$

where $R_N(k)$ is defined as

$$R_N(k) := \frac{1}{N}\sum_{t=1}^{N} \varepsilon_t \varepsilon_{t-k}$$

We note that $R_N(k)$ has the meaning of autocorrelation function of the prediction error.

Equation (3) is a nonlinear algebraic equation with respect to the parameter. In this case, the Newton-Rapson method [8] is used. Then the estimate is given by

$$\hat{\theta}^{i+1} = \hat{\theta}^{i} - \left[\frac{dR_N}{d\theta}\right]^{-1} R_N$$

where $\hat{\theta}^i$ is the parameter estimate at i'th iteration and R_N is defined by

$$R_N := \left[R_N(1) \quad \cdots \quad R_N(\quad + \quad)\right]^T \quad (4)$$

The estimates resulted by (4) may not have zeros inside the unit circle. We use the projection algorithm for this case.

Now, let's investigate the properties of DWM. The properties include the uniqueness, consistency, and asymptotic normality. For further statements, let's define as the compact subset of $R^{\ +}$ and $\overline{R}(k)$ as the time average of $\varepsilon_t\varepsilon_{t-k}$, that is,

$$\overline{R}(k) := \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} \varepsilon_t\varepsilon_{t-k}$$

Before deriving the main properties, we need to derive the convergence of $R_N(k)$ to $\overline{R}(k)$.

Proposition 3.1

With the assumptions on true model and projection operation on the estimated model,

$$\sup_{\theta\in} \left|R_N(k) - \overline{R}_N(k)\right| \to \infty, \quad . \quad . 1$$

$$as \quad N\to\infty \ for \quad k = 1, \cdots, \quad +$$

proof:

We assumed e_t is white Gaussian noise, which implies that e_t is i.i.d with zero mean and bounded fourth moment, and the model structure $(z)/A(z)$ is uniformly stable. Moreover, $\hat{A}(z)/\hat{\ }(z)$ is also a uniformly stable filter. These fulfills the assumptions of Lemma 8.4 in [7], which give the result of this proposition.

The derivation of the uniqueness use the residue integral form in calculation of $\overline{R}(k)$. From the relation

$$\varepsilon_t = \frac{\hat{A}(z) \quad (z)}{\hat{\ }(z)A(z)} e_t$$

, $\overline{R}(k)$ can be represented as

$$\overline{R}(k) = \frac{\sigma^2}{2\pi i}\oint \frac{\hat{A}(z) \quad (z)\hat{A}(z^{-1}) \quad (z^{-1})}{\hat{\ }(z)A(z)\ \hat{\ }(z^{-1})A(z^{-1})} z^k \frac{dz}{z}$$

where $\oint$ means the residue integral.

Following proposition shows that the estimate which make $\overline{R}(k)$ is unique and is the true value.

Proposition 3.2

If the estimates satisfy $\overline{R}(k)$=0 for $k = 1, \cdots, \quad +$, then

$$\hat{A}(z) = A(z) \quad and \quad \hat{\ }(z) = \quad (z)$$

Proof:

From lemma 1 in [10], $\hat{A}(z) \quad (z)\hat{A}(z^{-1}) \quad (z^{-1}) / \hat{\ }(z)A(z)\hat{\ }(z^{-1})A(z^{-1})$ is analytic in unit circle.

Since, $A(z)$, (z), $\hat{A}(z)$ and $\hat{\ }(z)$ have all zeros inside the unit circle and $A(z)$ and (z) have no common factors, it is clear that $\hat{A}(z) = A(z)$ and $\hat{\ }(z) = \quad (z)$.

From the convergence of $R_N(k)$ and the uniqueness of the estimate, the consistency is acquired as follows.

Proposition 3.3

If the estimates satisfy (3), then

$$\hat{\theta} \to \theta^* \quad . \quad . \quad as \quad N\to\infty$$

proof:

From the Proposition 3.1 and the compactness of , we get

$\hat{\theta} \to \ _T$, w. p. 1, as $N\to\infty$

where

$$_T = \left\{\theta \mid \theta\in \quad , \overline{R}(k) = 0,\ k = 1, \cdots, \quad + \quad\right\}$$

From the uniqueness of the estimate, the element of $_T$ is only θ^*, which completes the proof.

Now let's investigate the asymptotic normality. The asymptotic normality indicate how fast the estimate error decays as the number of samples used in estimation increases.

Proposition 3.4

Let's define $\overline{R} := \left[\overline{R}(1) \quad \cdots \quad \overline{R}(\quad + \quad)\right]^T$ and $R_N := \left[R_N(1) \quad \cdots \quad R_N(\quad + \quad)\right]^T$. Assume that $\left[d\overline{R}/d\theta\right]_{\theta=\theta^*}$ exists and is nonsingular and that

$\sqrt{N}\ \ [R_N]_{\theta=\theta^*} \rightarrow 0$ *as* $N \rightarrow \infty$. And also assume that the estimates satisfy (3), then $\sqrt{N}\left(\hat{\theta}-\theta^*\right)$ has normal distribution with mean zero and covariance P_θ *as* $N \rightarrow \infty$ *here*

$$P_\theta = \left[\frac{d\overline{R}}{d\theta}\right]^{-1}_{\theta=\theta^*} Q \left[\frac{d\overline{R}}{d\theta}\right]^{-T}_{\theta=\theta^*} \quad ith$$

$$Q = \lim_{N\rightarrow\infty} N\ \ \lfloor R_N^T R_N \rfloor_{\theta=\theta^*}$$

Proof:

We showed $\hat{\theta} \rightarrow \theta^*$ w. p. 1 as $N \rightarrow \infty$ at the Proposition 3.3. Moreover, by the projection of zeros of $\hat{A}(z)$ and $\hat{\ }(z)$ into unit circle, the $\hat{A}(z)/\ \hat{\ }(z)$ is uniformly stable and, therefore, its derivative with respect to θ is also uniformly stable. These fulfill the assumptions of Theorem 9.2 in [7]. Therefore, we get the results of this proposition.

From this proposition, we can say that $\hat{\theta}-\theta^*$ decays with covariance P_θ / N under some conditions : $\left[d\overline{R}/d\theta\right]_{\theta=\theta^*}$ exists and is nonsingular and that $\sqrt{N}\ \ [R_N]_{\theta=\theta^*} \rightarrow 0$ *as* $N \rightarrow \infty$.

4. SIMULATION

In this section, the performance of DWM is verified through the simulation. For simulation, following three ARMA(2,2) models used in [6] are considered.

Model S1:

$$(1+0.9z^{-1}+0.95z^{-2})\ \ _t = (1+1.25z^{-1}+0.75z^{-2})e_t$$

Model S2:

$$(1+0.9z^{-1}+0.95z^{-2})\ \ _t = (1+0.9z^{-1}+0.65z^{-2})e_t$$

Model S3:

$$(1-1.5z^{-1}+0.7z^{-2})\ \ _t = (1-0.7z^{-1}+0.25z^{-2})e_t$$

The models selected represent various systems which have different characteristics each other. The pole zero plots of the models are shown in Fig. 1. Model S1 and S2 have pole and zeros near the unit circle and according to the [6], the standard PLR method does not give accurate estimates for these two models. The algorithms given by (3) for PLR and (4) for DWM are applied to estimate the parameters. The obtained results are listed in Table 1. The standard PLR method does not converge for model S1 ~ S2 and converges closely to the true values for model S3, which coincide with the results in [6]. On the contrary DWM give the estimates close to true values for all models chosen.

5. CONCLUSION

We proposed a new variant of correlation approach for ARMA model. It is based on the idea to find the estimate making the prediction error at current time become uncorrelated with the one at past time. As good properties, we proved the uniqueness, consistency, and asymptotic normality of the estimate. With some simulations, we showed that the proposed method give good estimates for various systems compared to the standard PLR method.

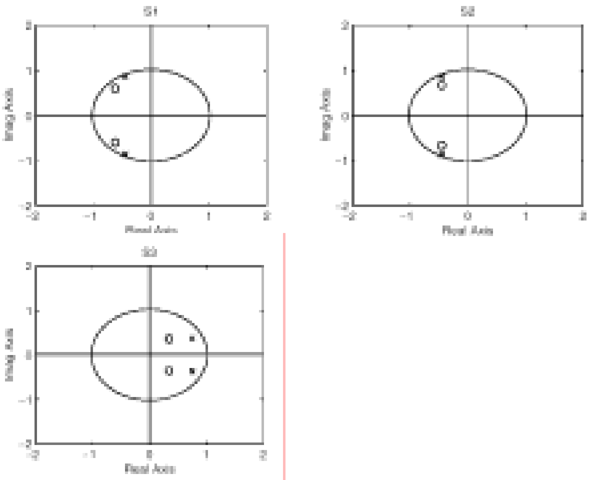

Fig. 1 The pole zero plot of the models chosen

Table 1 The estimation results

model	method	parameters			
		a1	a2	c1	c2
S1	true	0.9	0.95	1.25	0.75
	proposed	0.92	0.97	1.26	0.77
	PLR	1.15	1.0	1.19	0.49
S2	true	0.9	0.95	0.9	0.65
	proposed	0.91	0.98	0.91	0.68
	PLR	1.11	1.0	1.08	0.69
S3	true	-1.5	0.7	-0.7	0.25
	proposed	-1.47	0.67	-0.67	0.25
	PLR	-1.49	0.69	-0.69	0.25

REFERENCES

[1] Sоderstrоm, T. and Stoica, P., "Comparison of Some Instrumental Variable Methods - Consistency and Accuracy Aspects", *Auto atica* , 101-115, 1981

[2] Friedlander, B., 'The Overdetermined Instrumental Variable Method", *I Trans. on Auto atic ontrol A* , 353--356, 1984

[3] Porat, B. and Friedlander, B., "The Square-Root Overdetermined Instrumental Variable Algorithm", *I Trans. on Auto atic ontrol A* , 656--658, 1989

[4] Zhang, Y., Wen, C. and Soh, Y. C., "Indirect Closed-Loop Identification by Optimal Instrumental Variable Method," *Auto atica* , 2269-2271, 1997

[5] Stoica, P., Sоderstrоm, T., Ahlen, A. and Solbrand, G., "On the Asymptotic Accuracy of Pseudo-Linear Regression Algorithms", *Int. J. ontrol* , 115-126, 1984

[6] Stoica, P., Sоderstrоm, T., Ahlen, A. and Solbrand, G., "On the Convergence of Pseudo-Linear Regression Alogorithm", *Int. J. ontrol* , 1429-1444, 1985

[7] Ljung, L., System Identification Theory for the User, Prentice Hall, 1999

[8] Luenberger, D. G., *ti ization b ector ace ethods*, John Wiley and Sons, 1969

[9] Ljung, L., and Sderstrm, T., *Theor and Practice of Recursive dentification*, MIT press, 1987

[10] Ǻstrоm, K. J. and Sоderstrоm, T., "Uniquenness of the Maximun Likelihood Estimates of the Parameters of an ARMA Model", *I Trans. on Auto atic ontrol A* , 769-773, 1974

[11] Landau, I. D., *ste Identification and ontrol esign*, Prentice Hall, 1990