2G2-2

ANALYZING RELATIONSHIPS BETWEEN CTARMA AND ARMA MODELS

Marina Demeshko, Takashi Washio, Yoshinobu Kawahara and Shohei Shimizu

The Institute of Scientific and Industrial Research, Osaka University

Abstract

A linear Markov system can be represented by an autoregressive and moving average (ARMA) model in discrete time domain. It can be used to identify some system model and its associated parameters. Recently, the ARMA model has been extended to an ARMA-LiNGAM model which is a canonical form to represent the system. It is expected to provide more detailed information of the model structure and the parameters. In this paper, we extend the model to a more generic ARMA-LiNGM model and analyze the relationships between the ARMA-LiNGM model and a CTARMA model which is another canonical form of the system model in continuous time domain. As the consequence, we provide the relations between the coefficients of the two models, which can help us to overcome limitations of a classical ARMA model on the identification of the model and its parameters.

1. Introduction

Most of real world systems are characterized by their mathematical models and their parameters which provide their useful process information for various engineering purposes, but these models and parameter values are unknown in many cases. A good example is model identification of a nuclear power plant from measurement data. Most important process of a nuclear reactor is neutron breeding usually represented by a point-kinematics reactor model [Hetrick 71]:

$$\frac{dn(t)}{dt} = \frac{p-\beta}{l} n(t) + \sum_i \lambda_i c_i(t) + q, \qquad (1)$$

$$\frac{dc_i(t)}{dt} = \frac{\beta_i}{l} n(t) - \lambda_i c_i(t) \quad (i = 1 \dots 6), \qquad (2)$$

Contact: Demeshko M., ISIR Osaka University, 大阪吹田市津雲台 3-3、号 1319, 080-4246-0507, kaledom@gmail.com

where $n(t)$ and $c_i(t)$ are neutron flux and the i-th neutron emitter density at time t respectively. $ß_i$ and λ_i are respectively a delayed-neutron fraction and a decay constant of the i-th neutron emitter such that $\beta = \sum_i \beta_i$. p, l and q are reactivity, neutron generation time and neutron source intensity.

One of the biggest problem of nuclear reactor safety is to find up the model and its parameters p/l, β_i/l, λ_i and q. By knowing them, we can more effectively and safely control the reactor. But they are not directly obtained because these relations among the state variables are intentional. Accordingly, we need to indirectly estimate its structure and their values by using the directly measurable variables such as neutron flux n(t) only.

In case of a linear Markov system such as Eq.(1) and (2), we can apply ARMA modeling which can be used as an empirical system model obtained from the observed signals [Box 76]. But, the ARMA model has some limitations to identify our objective system model and its parameters. These limitations are: it does not give us the unique representation of the system because it is not canonical [Moneta 10], and it misses some instantaneous interactions among variables when the available granularity of the time discretization in the observation is not sufficient to capture the system dynamics. These limitations cause the difficulty to identify the model and its parameters of the system dynamics correctly.

Accordingly, we need to provide some extension of the ARMA model to include the instantaneous effects among variables and to propose its canonical form. Furthermore, if the canonical ARMA model is directly related with the differential equation model of the system, we can relate the empirical system analysis and the domain knowledge based system analysis. This enables the empirical validation of the domain knowledge on the system and the complement of the domain knowledge such as physical parameter value estimation. The research objective of this paper is to propose an appropriate canonical extension of the ARMA model named ARMA-LiNGM in discrete time domain which directly corresponds to some linear differential equation model of the system in continuous time domain and to give their quantitative relation. This work will provide the basis of the above mentioned unified framework of the empirical and deductive analyses.

2. Research Background

2-1. CTARMA model

Let`s consider a linear Markov system in continuous time domain. It can be represented by higher order linear differential equations. Their standard formulation is p-th ordinary linear differential equations and observation equations of the state space model in continuous time domain [Stamer 96]:

$$dX(t)/dt=AX(t)+BdW(t)/dt \quad (3),$$

$$Y(t)=C^tX(t) \quad (4)$$

where (3) is a Langevin differential equation for X(t) which is a vector of derivatives of a state variable vector x(t) as

$$X(t)=\begin{bmatrix} x(t) \\ \dots \\ x^{(p-1)}(t) \end{bmatrix},$$

where $x^{(i)}(t)$ is the i-th derivative of x(t). Here, A is a matrix having an upper shift block structure except the bottom p blocks S_m (m=0,…,p-1) representing system dynamics. B is a matrix having all zeros except an identity submatrix at its bottom block. C is a matrix having an identity submatrix at its top block, observation coefficient submatrices R_m (m=1,…,q) from the second to (q+1)-th blocks and zeros in the rest.

This model is easily transformed to the following CTARMA model in continuous time domain [Stamer 96]:

$$Y^{(p)}(t)=\sum_{m=0}^{p-1} S_m Y^{(m)}(t)+W^{(0)}(t)+\sum_{m=1}^{q} R_m W^{(m)}(t) \quad (5)$$

If we know a relation between a CTARMA model and an ARMA model empirically obtained from observed time series of Y(t), we immediately obtain the state space model: Eq.(3) and (4) because the coefficient matrices of the CTARMA model directly correspond to the matrices of the state space model. Thus, CTARMA model provides a connection between the empirical ARMA model and the dynamics model of a linear Markov system.

2-2. ARMA model and ARMA-LiNGAM model

An ARMA model is a representation of a linear Markov system approximated in discrete time domain. A multivariate ARMA model of order (p,q) is defined by [Box 76]:

$$Y(t)=\sum_{j=1}^{p}\Phi_j Y(t-j\Delta t)+U(t)+\sum_{j=1}^{q}\theta_j U(t-j\Delta t), \quad (6)$$

where U(t) is an external noise vector, and Φ_j, θ_j are AR and MA coefficient matrics. This is equivalent to

$$P^{-1}Y(t)=\sum_{j=1}^{p}P^{-1}\Phi_j P P^{-1}Y(t-j\Delta t)+U^*(t)$$

$$+\sum_{j=1}^{q}P^{-1}\theta_j P U^*(t-j\Delta t), \quad (7)$$

where P is a regular matrix and $U^*(t)=P^{-1}U(t)$. This implies that a given impulse response of the ARMA model Eq.(6):

$$Y(t)=\sum_{j=0}^{\infty}\Phi_j U(t-j)$$

is equivalent to

$$Y(t)=\sum_{j=0}^{\infty}\Phi_j P P^{-1}U(t-j)=\sum_{j=0}^{\infty}\Phi_j P U^*(t-j)$$

Thus Y(t) has a same response for different U(t) and $U^*(t)$ under P. It is clear that the representation of the ARMA model is not unique, because for any nonsingular matrix P we will get equivalent impulse responses.

That can be solved by using the linear non-Gaussian acyclic model (ARMA-LINGAM), which is a combination of autoregressive and structural-equation models: $Y(t) = B_0 Y(t) + \sum_{\tau=1}^{k} B_\tau Y(t-\tau) + W(t)$, where the LiNGAM part is acyclic (B_0 can be represented by a lower triangular matrix), and each element of $W(t)$ is a non-Gaussian noise. These assumptions of acyclicity and non-Gaussianity enables the estimation of the LiNGAM part: $B_0 x(t)$ of the model [Hyvärinen 08]. By this LiNGAM part, ARMA-LINGAM represents instantaneous influences among variables. But it assumes acyclic dependency of the instantaneous influences among variables which does not generally hold for most physical processes.

3. Proposed ARMA-LiNGM model

A solution to the above limitation is to retract the acyclicity assumption. Instead, we assume that the process of the objective system follows a linear Markov system represented by a linear differential equations in continuous time domain, and we propose an extension of an ARMA-LiNGAM model named "auto-regressive moving average and linear non-Gaussian model (ARMA-LiNGM)". :

$$Y(t) = \Psi_0 Y(t) + \sum_{m=1}^{p} \Psi_m Y(t - m\Delta t) + W(t) + \sum_{m=1}^{q} \Omega_m W(t - m\Delta t) \quad \text{, or}$$

$$(I - \Psi_0) Y(t) = \sum_{m=1}^{p} \Psi_m Y(t - m\Delta t) + W(t) + \sum_{m=1}^{q} \Omega_m W(t - m\Delta t) \quad (8)$$

Here, Ψ_0 is an arbitary matrix, but its diagonals are zero, and $I - \Psi_0$ is regular. $W(t)$ is a non-Gaussian noise vector. This is the unique difference from Ψ_0 of the ARMA-LiNGAM.

If we compare this model Eq.(8) with the ARMA model Eq.(6), we understand that an ARMA-LINGM is a representation of the ARMA model, and their coefficient matrices have the following relations [Kawahara 10]:

$$\Psi_j = (I - \Psi_0)^{-1} \Phi_j$$

$$W(t) = (I - \Psi_0)^{-1} \varepsilon(t) \quad (9)$$

$$\Omega_j = (I - \Psi_0) \theta_j (I - \Psi_0)^{-1}$$

We understand that ARMA-LINGM is canonical under a given Ψ_0 by comparing Eq.(7) and (8) since the matrix P is uniquely specified by $P = I - \Psi_0$.

4. Relation of CTARMA and ARMA-LiNGM

Our objective in this paper is to analyze relations of coefficient matrices between a CTARMA model and our proposed ARMA-LiNGM by assuming that both models represent an identical linear Markov system under a sampling time discretization. As shown later, the consequence of this analysis provides a strong constraint to determine Ψ_0.

First, we introduce an approximation of CTARMA model for discrete time. This is made through discrete approximation of n-th derivative of Y(t) using Euler approximation:

$$\frac{d^n Y(t)}{dt^n} = \frac{Y_{n-1}(t) - Y_{n-1}(t - \Delta t)}{\Delta t}$$

It is convenient to apply here Leibniz rule [Olver 93]:

$$Y_n(t) = \frac{1}{\Delta t^n} \sum_{k=0}^{n} (-1)^k \frac{n!}{(n-k)!k!} Y(t - k\Delta t) \quad (10)$$

By substituting Eq.(10) into Eq.(5), we obtain the next form:

$$\left(\sum_{k=0}^{q} R_k \Delta t^{p-k}\right)^{-1} \left(I - \sum_{m=0}^{p-1} \Delta t^{p-m} S_m\right) Y(t) W(t - m\Delta t),$$

$$= \left(\sum_{k=0}^{q} R_k \Delta t^{p-k}\right)^{-1} \sum_{m=1}^{p} (-1)^m \sum_{k=m}^{p} \frac{k!}{(k-m)!\,m!} S_k \Delta t^{p-k} Y(t - m\Delta t) + W(t) +$$

$$+ \left(\sum_{k=0}^{q} R_k \Delta t^{p-k}\right)^{-1} \sum_{m=1}^{q} (-1)^m \sum_{k=m}^{q} \frac{k!}{(k-m)!\,m!} R_k \Delta t^{p-k}$$

$$\text{where } S_p = -I, R_0 = I\Delta t^{-p}. \quad (11)$$

By comparing Eq.(8) and Eq.(11), we obtain the following Lemma.

Lemma 1 Coefficient matrices of an ARMA-LiNGM in Eq.(8) is represented by coefficient matrices of a

CTARMA model in Eq.(5) as follows.

$$\Psi_0 = I - \left(\sum_{k=0}^{q} R_k \Delta t^{p-k}\right)^{-1} \left(I - \sum_{m=0}^{p-1} \Delta t^{p-m} S_m\right),$$

$$\Psi_m = (-1)^m \left(\sum_{k=0}^{q} R_k \Delta t^{p-k}\right)^{-1} \sum_{k=m}^{p} \frac{k!}{(k-m)!\,m!} S_k \Delta t^{p-k}$$

$$\Omega_m = (-1)^m \left(\sum_{k=0}^{q} R_k \Delta t^{p-k}\right)^{-1} \sum_{k=m}^{q} \frac{k!}{(k-m)!\,m!} R_k \Delta t^{p-k}$$

where $m \geq 1, S_p = -I, R_0 = I\Delta t^{-p}$.

When a CTARMA model is given, then we uniquely obtain an ARMA-LiNGM by this Lemma.

In reverse, we can get the coefficient matrices of a CTARMA model: R_m and S_m by an ARMA-LiNGM derived from a given time series data of Y(t). It is done by solving the last two equations in Lemma 1. Considering $R_0 = I\Delta t^{-p}$ and $S_p = -I$, we can solve the formulae with R_k and S_k by an inductive derivation and obtain the next Lemma.

Lemma 2 Coefficient matrices of a CTARMA model in Eq.(5) is represented by coefficient matrices of an ARMA-LiNGM in Eq.(8) as follows.

$$R_m = (-1)^m \Delta t^{m-p} \left(I - \sum_{s=1}^{q} (-1)^s \sum_{k=s}^{q} \frac{k!}{(k-s)!\,s!} \Omega_k\right)^{-1} \sum_{k=m}^{q} \frac{k!}{(k-m)!\,m!} \Omega_k$$

, where $1 \leq m \leq q$

$$S_n = \Delta t^{n-p} (-1)^n \left(\left(I - \sum_{s=1}^{q} (-1)^s \sum_{k=s}^{q} \frac{k!}{(k-s)!\,s!} \Omega_k\right)^{-1} \sum_{k=n}^{p-1} \frac{k!}{(k-n)!\,n!} \Psi_k + (-1)^{p-1} \frac{p!}{(p-n)!\,n!} I\right)$$

, where $1 \leq n \leq p-1$

Conclusion

We showed that there is a non-linear relation between the coefficients of the two models, CTARMA and ARMA-LINGM, upon a discrete time approximation of CTARMA. Accordingly, we can estimate a structure and parameters of a linear differential equations representing an objective system from an ARMA-LINGM model obtained from a given time series Y(t).

References

[Hetrick 71] David L. Hetrick.: Dynamics of nuclear reactor. *The University of Chicago Press, 3-6*. (1971)

[Box 76] Box G. and Jenkins G.M. Time Series Analysis: Forecasting and Control, second edition. Oakland, CA: Holden-Day (1976)

[Moneta 10] Moneta A., Entner D., Hoyer P., Coad A.: Causal Inference by Independent Component Analysis with Applications to Micro- and Macroeconomic Data. *Jena Economic Research Papers #2010-031, pp.3-4.* (2010)

[Stamer 96] Stamer O., Tweedie R.L., Brockwell P.J.: Existence and stability of continuous time threshold ARMA process. *Statistica Sinica, 6*, *pp.* 715-732 (1996)

[Hyvärinen 08] Hyvärinen A., Shimizu S., Hoyer P.: Casual Modelling Combining Instantaneous and Lagged Effects: an Identifiable Model Based on Non-Gaussianity. *Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland.* (2008)

[Kawahara 10] Kawahara Y., Shimizu S., Washio T.: Analyzing relationships between ARMA Processes Based on Non-Gaussianity of External Influences. *Preprint submitted to Neurocomputing* (2010)

[Olver 93] Olver P.: Applications of Lie groups to differential equations. *New-York, page 318*. (1993)