Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

The gradient algorithm for parameter and output estimation for dual-rate CARARMA systems ⋆

Huizhong Yang. * **Jun Tian** *

* *School of Communication and Control Engineering, Jiangnan University, 1800 Lihu Rd.,Wuxi, Jiangsu 214122, P.R.China (e-mail: yhz@jiangnan.edu.cn).*

Abstract: A recursive generalized extended stochastic forgetting gradient algorithm is used to identify the dual-rate stochastic systems based on the polynomial transformation technique. A time-varying forgetting factor is included to improve the rate of convergence. The intersample output estimation algorithm is also studied in the paper. Finally, a simulation example shows that the algorithm is excellent effective in parameter identification and output estimation.

1. INTRODUCTION

This paper deals with a dual-rate systems where the output $y(kJT)$ is sampled at a slower rate than the input $u(kT)$ and the output is needed at the fast rate. The least squares algorithms are used to estimate the parameters for dual-rate systems (Gudi *et al.* 1995; Huang *et al.* 1993; Lu and Fisher 1988,1989; Ding and Chen 2005a). But these algorithms mainly deal with the noise-free systems. In multirate stochastic systems with noise, Ding and Chen al. have proposed least squares algorithms for dual-rate systems and intersample estimation for the dual-rate CAR model, especially the convergence properties of the algorithms (Ding and Chen 2004a,2004b,2004c,2005b). Ding and Chen (2002,2007) have also studied the convergence properties of the stochastic gradient algorithm for CAR and CARMA models in single-rate system. Compared with the least squares, the forgetting gradient algorithm has low computation burden and the convergence rate of the parameter estimation can improve due to the presence of forgetting factor λ. So Yang and Tian (2006)has studied the dual-rate stochastic forgetting gradient algorithm for the CAR model. But for more complicated dual-rate CARARMA model, parameter identification and output estimation haven't been studied so for. The main purpose of this paper is parameter identification and output estimation for the dual-rate CARARMA system.

2. THE DESCRIPTION OF THE ALGORITHM

Assume that linear discrete time-invariant single-rate process model is represented by the following equation,

$$a(z)y(kT) = b(z)u(kT), \tag{1}$$

where

$$a(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + ... + a_n z^{-n} = \prod_{i=1}^{n}[1 - \lambda_i z^{-1}],$$

$$b(z) = b_1 z^{-1} + b_2 z^{-2} + ... + b_n z^{-n}.$$

$u(kT)$ and $y(kT)$ are the input and output sequences of the system, respectively. The available measurement data are $u(kT)$ and $y(kJT)$, and the $[y(kJT + iT), i = 1, ..., J - 1]$ are not available due to the sensor constraints in output sampling. Then the polynomial transformation technique in (Ding and Chen 2005c) is used to transform the model (1) into a form which can be identified from the available dual-rate data.

Define the following polynomial equation

$$\phi_J(z) = \prod_{i=1}^{n}(1 + \lambda_i z^{-1} + \lambda_i^2 z^{-2} + ... + \lambda_i^{J-1} z^{-J+1}).$$

Multiplying both sides of equation(1) by $\phi_J(z)$ gets a new model:

$$\alpha(z^J)y(kJ) = \beta(z)u(kJ), \tag{2}$$

where

$$\alpha(z^J) = \phi_J(z)a(z) = 1 + \alpha_1 z^{-J} + \alpha_2 z^{-2J} + ... + \alpha_n z^{-nJ},$$
$$\beta(z) = \phi_J(z)b(z) = \beta_1 z^{-1} + \beta_2 z^{-2} + ... + \beta_{nJ} z^{-nJ}.$$

Introduce Autoregressive Moving average noise model Equation (2) can be expressed as

$$\alpha(z^J)y(kJ) = \beta(z)u(kJ) + \frac{d(z)}{c(z)}v(kJ). \tag{3}$$

where

$$c(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + ... + c_{n_c} z^{-n_c},$$
$$d(z) = 1 + d_1 z^{-1} + d_2 z^{-2} + ... + d_{n_d} z^{-n_d}.$$

Defining

$$e(kJ) = \frac{d(z)}{c(z)}v(kJ). \tag{4}$$

Equation (3) and (4) can be written as the recursive equation,

$$y(kJ) = \varphi_s^T(kJ)\theta_s + e(kJ), \tag{5}$$
$$e(kJ) = \varphi_n^T(kJ)\theta_n + v(kJ). \tag{6}$$

where

$$\theta_s^T(kJ) := [\alpha_1(kJ), \alpha_2(kJ), ..., \alpha_n(kJ);$$
$$\beta_1(kJ), \beta_2(kJ), ..., \beta_{nJ}(kJ)],$$

⋆ This research was supported by the the National Natural Science Foundation of China (No. 60674092) and the high-tech research project of Jiangsu province(No.BG200610).

$$\begin{aligned}\theta_n^T(kJ) :=& [c_1(kJ), c_2(kJ), ..., c_{n_c}(kJ);\\ & d_1(kJ), d_2(kJ), ..., d_{n_d}(kJ)],\\ \varphi_s^T :=& [-y(kJ-J), -y(kJ-2J)..., -y(kJ-nJ);\\ & u(kJ-1), u(kJ-2), ..., u(kJ-nJ)],\\ \varphi_n^T :=& [-e(kJ-1), -e(kJ-2)..., -e(kJ-n_c);\\ & v(kJ-1), v(kJ-2), ..., v(kJ-n_d)].\end{aligned}$$

To simplify the computation process, (5) and (6) can be combined into the following equation,

$$y(kJ) = \varphi^T(kJ)\theta + v(kJ), \tag{7}$$

where

$$\begin{aligned}\theta^T(kJ) &= [\theta_s^T(kJ), \theta_n^T(kJ)],\\ \varphi^T(kJ) &= [\varphi_s^T(kJ), \varphi_n^T(kJ)].\end{aligned}$$

The recursive extended identification algorithm can be used to identify this dual-rate CARARMA model by increasing the dimensions of the parameter vector and information vector, adding the parameters c_i and d_i of the noise model to the parameter vector and adding the noise terms $e(kJ-i), i = 1, ..., n_c$ and $v(kJ-i), i = 1, ..., n_d$ to the information vector. Since $e(kq-i)$ and $v(kq-i)$ are unknown, they are replaced by the estimation $\hat{e}(kq-i)$ and $\hat{v}(kq-i)$,

$$\hat{e}(kJ-i) = y(kJ-i) - \varphi_s^T(kJ-i)\hat{\theta}_s(kJ), i = 1, ..., n_c, \tag{8}$$

$$\hat{v}(kJ-i) = y(kJ-i) - \varphi^T(kJ-i)\hat{\theta}(kJ), i = 1, ..., n_d, \tag{9}$$

or,

$$\hat{v}(kJ-i) = \hat{e}(kJ-i) - \varphi_n^T(kJ-i)\hat{\theta}_n(kJ), i = 1, ..., n_d.$$

But actually the right-hand sides of (8) and (9) contain the unknown intersample outputs $y(kJ-i)$. Thus, it is impossible to compute the $\hat{e}(kJ-i)$ and $\hat{v}(kJ-i)$ by (8) and (9). A feasible way is that after computing the parameter estimate $\hat{\theta}$ at each step, $y(kJ-i)$ is replaced by its estimation $\hat{y}(kJ-i)$. Ding and Chen has given a method to determine the single-rate model to compute the $\hat{y}(kJ-i)$ (Ding and Chen 2005a). But a more simple way is to compute the $\hat{y}(kJ-i)$ directly from the following equation,

$$\hat{y}(kJ+i) = \begin{cases} y(kJ)\ , & i = 0,\\ \hat{\varphi_s}^T(kJ+i)\hat{\theta}_s(kJ), & i = 1, \ldots, J-1, \end{cases} \tag{10}$$

where

$$\begin{aligned}\hat{\varphi}_s(kJ+i) = [&-\hat{y}(kJ-J+i), -\hat{y}(kJ-2J+i),\\ &\cdots, -\hat{y}(kJ-nJ+i)\\ &u(kJ+i), u(kJ+i-1),\\ &\cdots, u(kJ+i-nJ)]^T.\end{aligned}$$

Then, the forgetting factor gradient algorithm of estimating the parameter vector of the dual-rate stochastic systems in (3) may be expressed as dual-rate recursive generalized extended forgetting gradient algorithm (DR-RGEFG algorithm for short):

$$\begin{aligned}\hat{\theta}(kJ) &= \hat{\theta}(kJ-J) + \frac{\hat{\varphi}(kJ)}{r(kJ)}\\ &\times [y(kJ) - \hat{\varphi}^T(kJ)\hat{\theta}(kJ-J)], \qquad (11)\\ \hat{\theta}(kJ+i) &= \hat{\theta}(kJ), i = 1, \cdots, J-1, \qquad (12)\\ r(kJ) &= \lambda r(kJ-J) + \|\hat{\varphi}(kJ)\|^2,\\ r(0) &= 1,\ 0 \le \lambda \le 1, \qquad (13)\\ \hat{\theta}^T(kJ) &= [\hat{\theta}_s^T(kJ), \hat{\theta}_n^T(kJ)],\\ \hat{\varphi}^T(kJ) &= [\hat{\varphi}_s^T(kJ), \hat{\varphi}_n^T(kJ)],\end{aligned}$$

$$\begin{aligned}\theta_s^T(kJ) &= [\alpha_1(kJ), \alpha_2(kJ), ..., \alpha_n(kJ);\\ &\quad \beta_1(kJ), \beta_2(kJ), ..., \beta_{nJ}(kJ)],\\ \theta_n^T(kJ) &= [c_1(kJ), c_2(kJ), ..., c_{n_c}(kJ);\\ &\quad d_1(kJ), d_2(kJ), ..., d_{n_d}(kJ)],\end{aligned}$$

$$\begin{aligned}\varphi_s^T &= [-y(kJ-1), -y(kJ-2)..., -y(kJ-n);\\ &\quad u(kJ-1), u(kJ-2), ..., u(kJ-nJ)],\\ \hat{\varphi}_n^T &= [-\hat{e}(kJ-1), -\hat{e}(kJ-2)..., -\hat{e}(kJ-n_c);\\ &\quad \hat{v}(kJ-1), \hat{v}(kJ-2), ..., \hat{v}(kJ-n_d)],\end{aligned}$$

$$\hat{e}(kJ-i) = \hat{y}(kJ-i) - \varphi_s^T(kJ-i)\hat{\theta}_s(kJ), i = 1, ..., n_c,$$

$$\hat{v}(kJ-i) = \hat{y}(kJ-i) - \hat{\varphi}^T(kJ-i)\hat{\theta}(kJ), i = 1, ..., n_d,$$

or

$$\hat{v}(kJ-i) = \hat{e}(kJ-i) - \hat{\varphi}_n^T(kJ-i)\hat{\theta}_n(kJ), i = 1, ..., n_d.$$

3. EXAMPLES

Assume that discrete-time system model takes the following form,

$$\begin{aligned}p(z) &= \frac{b(z)}{a(z)},\\ a(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} = 1 - 1.5z^{-1} + 0.7z^{-2},\\ b(z) &= b_1 z^{-1} + b_2 z^{-2} = z^{-1} + 0.5z^{-2}.\end{aligned}$$

The corresponding dual-rate models with the noise can be expressed as

$$\alpha(z^J)y(kJ) = \beta(z)u(kJ) + \frac{d(z)}{c(z)}v(kJ),$$

where

$$\begin{aligned}\phi_J(z) &= 1 + 1.5z^{-1} + 0.7z^{-2},\\ c(z) &= 1 - z^{-1} + 0.41z^{-2}.\\ d(z) &= 1 + 0.1z^{-1}.\end{aligned}$$

Taking $J = 2$, i.e., $\{u(k), y(2k)\}$ are available data. Here $\{u(k)\}$ is taken as a persistent excitation sequence with zero mean and unit variance, and $\{v(k)\}$ as a white noise sequence with zero mean and variance σ_v^2.Applying the DR-RGEFG algorithm to estimate the parameters (α_i, β_i) of this system. The parameter estimates with different data length are shown in Table 1, where δ is the relative parameter estimation error measured in the Euclideannorm:$\delta = \|\hat{\theta}(kJ) - \theta\|/\|\theta\|$. θ is the real parameter value, $\hat{\theta}$ is the estimation of the θ.

Fig. 1 shows δ with different $\lambda's$. The smaller λ is, the faster parameter estimation converges. But a smaller λ results in large parameter estimation variance. Though the bigger λ results in slow convergence rate, the parameter estimation error becomes large. A feasible method is to change λ with different k. In this example, take $\lambda = 0.7 + \frac{0.3*k}{length}$, where length is the data length. When $k = 0, \lambda = 0.7$, and when $k = length, \lambda = 1$. Fig. 2 illustrates parameter estimation verse k and Fig. 3 shows the parameter estimation error δ when $\lambda = 0.7 + \frac{0.3*k}{length}$, it can be seen that the parameter estimation is smooth

and fast. Fig. 4 illustrates the simulation run for output estimation error(OEE) $y(k) - \hat{y}(k)$. $y(k)$ is system output and $\hat{y}(k)$ is the output estimation. In Fig. 3, the OEE becomes smaller when $k \to \infty$.

Table 1. The DR-RGEFG estimate of parameter($\sigma_v^2 = 1.00, \delta_{ns} = 63.695\%$)

λ	k	α_1	α_2	β_1
1	500	-0.63548	0.20294	0.10818
	1000	-0.66862	0.24597	0.12233
	2000	-0.69826	0.27932	0.13572
0.9	500	-0.94082	0.54884	0.74686
	1000	-0.89598	0.52241	0.93541
	2000	-0.86304	0.50453	0.99454
0.7	500	-0.87209	0.50621	0.99909
	1000	-0.84367	0.49993	1.00664
	2000	-0.84821	0.49346	1.00856
0.5	500	-0.86079	0.49886	1.00348
	1000	-0.83715	0.50045	1.00811
	2000	-0.84771	0.49164	1.01184
$\lambda = 0.7 + \frac{0.3*k}{length}$	500	-0.87447	0.50767	0.99508
	1000	-0.84877	0.49996	1.00629
	2000	-0.85003	0.49480	1.00210
	θ	-0.85000	0.49000	1.00000

λ	k	β_2	β_3	β_4	δ
1	500	0.44479	0.28311	0.14316	0.76142
	1000	0.46894	0.30332	0.15282	0.74634
	2000	0.49226	0.32334	0.16115	0.73247
0.9	500	1.54987	1.01610	0.21262	0.24352
	1000	1.87374	1.28193	0.23201	0.08941
	2000	1.99018	1.42972	0.31292	0.01674
0.7	500	1.98552	1.40821	0.30059	0.02508
	1000	1.99856	1.45006	0.34654	0.00490
	2000	1.99614	1.45673	0.35669	0.00486
0.5	500	2.00366	1.44614	0.35877	0.00618
	1000	2.00115	1.45636	0.34829	0.00686
	2000	1.99632	1.45479	0.35516	0.00508
$\lambda = 0.7$ $+\frac{0.3*k}{length}$	500	1.97891	1.39632	0.28843	0.03136
	1000	1.99536	1.44379	0.34140	0.00579
	2000	1.99887	1.45551	0.35574	0.00335
	θ	2.00000	1.45000	0.35000	

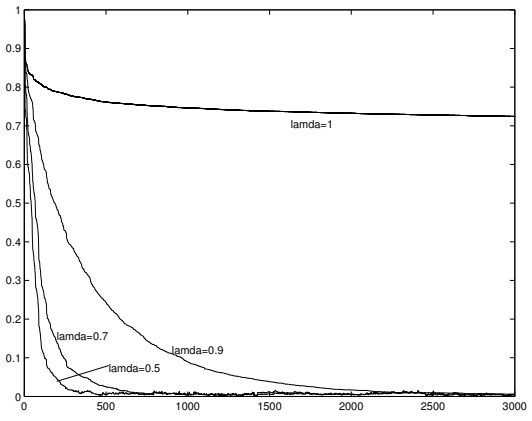

Fig. 1. Parameter estimation $\hat{\theta}$ verses k, ($\sigma_v^2 = 1.00^2, \delta_{ns} = 63.695\%$)

4. CONCLUSIONS

A dual-rate recursive generalized extended forgetting gradient algorithm is used to estimate the parameters of the dual-rate CARARMA systems which the input is sampled J times faster than the output; The algorithm uses only dual-rate measurement data. For the unknown noises $v(k)$ and $e(k)$, we replace $v(k)$ and $e(k)$ with $\hat{v}(k)$ and $\hat{e}(k)$ which applies the output estimation $\hat{y}(k)$. The result shows that parameter estimation error is small. Based on the estimated models, intersample output estimation is also studied. The example shows the effectiveness of the algorithm.

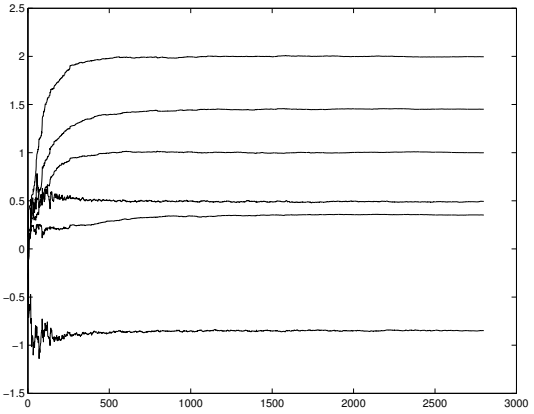

Fig. 2. Parameter estimation $\hat{\theta}$ verses k, ($\sigma_v^2 = 1.00^2, \delta_{ns} = 63.695\%$)

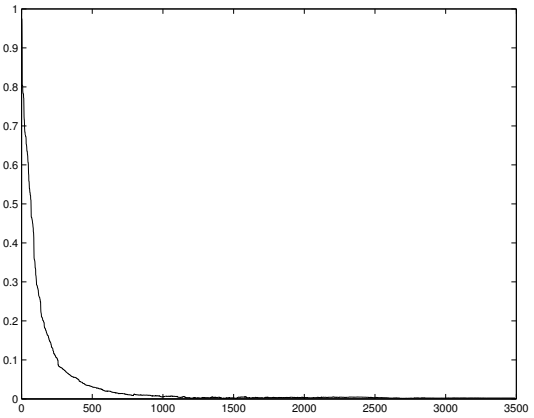

Fig. 3. The estimation errors δ verses k with $\lambda = 0.7 + \frac{0.3*k}{length}$, ($\sigma_v^2 = 1.00^2, \delta_{ns} = 63.695\%$)

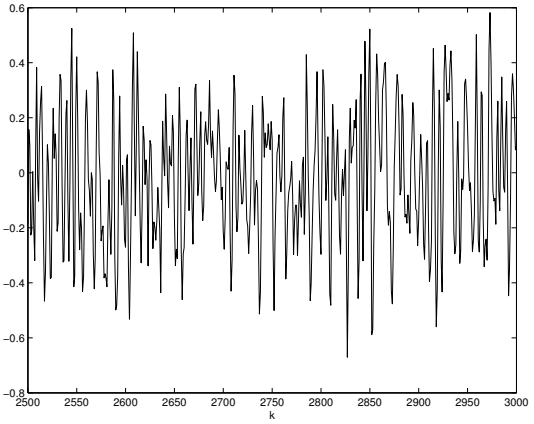

Fig. 4. The output estimation errors , $\lambda = 0.7 + \frac{0.3*k}{length}$, ($\sigma_v^2 = 1.00^2, \delta_{ns} = 63.695\%$)

REFERENCES

Gudi, R.D., Shah, S.L. and Gray, M.R. (1995). Adaptive multirate state and parameter estimation strategies with application to a bioreactor, *AIChE J*, **41**, 2451-2464.

Huang, Z.W., Xu, Y.M., Fang, C.Z. and Tang, J. (1993). Improvements in dynamic compartmental modelling of distillation columns, *Journal of Process Control*, **3**, 139-145.

Lu, W.P. and Fisher, D.G. (1988). Output estimation with multi-rate sampling, *International Journal of Control*, **48**, 149-160.

Lu, W.P. and Fisher, D.G. (1989). Least-squares output estimation with multirate sampling, *IEEE Transactions on Automatic Control*, **34**, 669-672.

Ding F. and Chen T. (2005a). Hierarchical identification of lifted state-space models for general dual-rate systems, *IEEE Transactions on Circuits and Systems–I*, **52**, 1179-1187.

Ding, F. and Chen, T. (2005b). Parameter estimation of dual-rate stochastic systems by using an output error method, *IEEE Transactions on Automatic Control*, **50**, 1436-1441.

Ding, F. and Chen, T. (2004a). Combined parameter and output estimation of dual-rate systems using an auxiliary model, *Automatica*, **40**, 1739-1748.

Ding, F. and Chen,T. (2004b). Least squares based self-tuning control of dual-rate systems, *Int. J. Adapt. Control Signal Process*, **18**, 697-714.

Ding, F. and Chen, T. (2004c). Identification of dual-rate systems based on finite impulse response models, *International Journal of Adaptive Control and Signal Processing*, **18**, 589-598.

Ding, F. and Ding, T. (2002). Convergence of forgetting gradient estimation algorithm for time-varying parameters, *Acta Automatica Sinica*, **28**, 962-968.

Ding F. and Chen T. (2007). Performance analysis of multi-innovation gradient type identification methods, *Automatica*, **43**, 1-14.

Yang, H.Z., Tian, J. and Ding, F. (2006). The forgetting gradient algorithm for parameter and intersample estimation of dual-rate systems, Neural Information Processing, PT 2, *Proceedings Lecture Notes In Computer Science*, **4233**, 721-728.

Ding, F. and Chen, T. (2005c). Modeling and identification for multirate systems, *Acta Automatica Sinica*, **31**, 105-122.