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 $\lambda$. 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 far. The main purpose of this paper is to 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,

$$a(z)y(kT) = b(z)u(kT),$$

where

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

and

$$b(z) = b_1 z^{-1} + b_2 z^{-2} + \ldots + b_n z^{-n}.$$
\[\theta_n^T(kJ) := [c_1(kJ), c_2(kJ), \ldots, c_{nc}(kJ); d_1(kJ), d_2(kJ), \ldots, d_{nd}(kJ)],\]
\[\varphi_s^T := [-y(kJ - J), -y(kJ - 2J), \ldots, -y(kJ - nJ); u(kJ - 1), u(kJ - 2), \ldots, u(kJ - nJ)],\]
\[\varphi_n^T := [-e(kJ - 1), -e(kJ - 2), \ldots, -e(kJ - n_c); v(kJ - 1), v(kJ - 2), \ldots, v(kJ - n_d)].\]

To simplify the computation process, (5) and (6) can be combined into the following equation,
\[y(kJ) = \varphi^T(kJ)\theta + v(kJ), \quad (7)\]
where
\[\theta^T(kJ) = [\theta_1^T(kJ), \theta_2^T(kJ)],\]
\[\varphi^T(kJ) = [\varphi_1^T(kJ), \varphi_2^T(kJ)].\]

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 \(e(kJ - i), i = 1, \ldots, n_c\) and \(v(kJ - i), i = 1, \ldots, n_d\) to the information vector. Since \(c(kJ - i)\) and \(v(kJ - i)\) are unknown, they are replaced by the estimation \(\hat{c}(kJ - i)\) and \(\hat{v}(kJ - i)\),
\[\hat{c}(kJ - i) = \hat{y}(kJ - i) - \varphi_s^T(kJ - i)\hat{\theta}_s(kJ), \quad i = 1, \ldots, n_c, \quad (8)\]
\[\hat{v}(kJ - i) = \hat{y}(kJ - i) - \varphi^T(kJ - i)\hat{\theta}(kJ), \quad i = 1, \ldots, n_d. \quad (9)\]

But actually the right-hand sides of (8) and (9) contain the unknown intersample outputs \(y(kJ - i)\). Thus, it is impossible to compute \(\hat{c}(kJ - i)\) 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,
\[\hat{y}(kJ + i) = \begin{cases} y(kJ), & i = 0, \\ \varphi_s^T(kJ + i)\hat{\theta}_s(kJ), & i = 1, \ldots, J, \end{cases} \quad (10)\]
where
\[\varphi_s(kJ + i) = [-\hat{y}(kJ - J + i), -\hat{y}(kJ - 2J + i), \ldots, -\hat{y}(kJ - n_J + i), \quad u(kJ + i), u(kJ + i - 1), \quad u(kJ + i - n_d)].\]

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 (DRRGEFG algorithm for short):
\[\hat{\theta}(kJ) = \hat{\theta}(kJ - J) + \frac{\varphi(kJ)}{r(kJ)}\]
\[\times [y(kJ) - \varphi^T(kJ)\hat{\theta}(kJ - J)], \quad (11)\]
\[\hat{\theta}(kJ + i) = \hat{\theta}(kJ), \quad i = 1, \ldots, J - 1, \quad (12)\]
\[r(kJ) = \lambda r(kJ - J) + \|\varphi(kJ)\|^2, \quad (13)\]

3. EXAMPLES

Assume that discrete-time system model takes the following form,
\[p(z) = \frac{b(z)}{a(z)} \]
\[a(z) = 1 + a_1z^{-1} + a_2z^{-2} = 1 - 1.5z^{-1} + 0.7z^{-2}, \quad b(z) = b_1z^{-1} + b_2z^{-2} = z^{-1} + 0.5z^{-2}.\]

The corresponding dual-rate models with the noise can be expressed as
\[\alpha(z^d)y(kJ) = \beta(z)u(kJ) + \frac{d(z)}{e(z)}r(kJ), \quad (14)\]
where
\[\phi_r(z) = 1 + 1.5z^{-1} + 0.7z^{-2}, \quad c(z) = 1 - z^{-1} + 0.41z^{-2}, \quad d(z) = 1 + 0.1z^{-1}.\]

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 \(\sigma_v^2\). Applying the DR-RGEFG algorithm to estimate the parameters \(\{\alpha_1, \beta_1\}\) of this system. The parameter estimates with different length are shown in Table 1, where \(\delta\) is the relative parameter estimation error measured in the Euclidean norm; \(\delta = \|\hat{\theta}(kJ) - \theta\|/\|\theta\|\). \(\hat{\theta}\) is the real parameter value, \(\hat{\theta}\) is the estimation of the \(\theta\). Fig. 1 shows \(\delta\) with different \(\lambda\)'s. The smaller \(\lambda\) is, the faster parameter estimation converges. But a smaller \(\lambda\) results in large parameter estimation variance. Though the bigger \(\lambda\) results in slow convergence rate, the parameter estimation error becomes large. A feasible method is to change \(\lambda\) with different \(k\). In this example, take \(\lambda = 0.7 + \frac{0.3}{\text{length}}\) where length is the data length. When \(k = 0, \lambda = 0.7, \text{ and when } k = \text{length}, \lambda = 1. \) Fig. 2 illustrates parameter estimation verse \(k\) and Fig. 3 shows the parameter estimation error \(\delta\) when \(\lambda = 0.7 + \frac{0.3}{\text{length}}\); 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($\sigma_v^2 = 1.00, \delta_{ns} = 63.695\%$)

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$k$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>-0.63548</td>
<td>0.20294</td>
<td>0.10818</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-0.66862</td>
<td>0.24397</td>
<td>0.12233</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.69826</td>
<td>0.27932</td>
<td>0.13572</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>500</td>
<td>-0.34582</td>
<td>0.54884</td>
<td>0.74686</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-0.89598</td>
<td>0.52241</td>
<td>0.93541</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.86304</td>
<td>0.50453</td>
<td>0.99454</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>500</td>
<td>-0.87209</td>
<td>0.50621</td>
<td>1.00210</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-0.84367</td>
<td>0.49993</td>
<td>1.00664</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.84821</td>
<td>0.49346</td>
<td>1.00856</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda = 0.7 + \frac{0.3 \times k}{v_{length}}$</td>
<td>1000</td>
<td>-0.86079</td>
<td>0.39886</td>
<td>1.00398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.83715</td>
<td>0.30045</td>
<td>1.00811</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda = 0.7 + \frac{0.3 \times k}{v_{length}}$</td>
<td>500</td>
<td>-0.87447</td>
<td>0.50767</td>
<td>0.99568</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-0.84877</td>
<td>0.49996</td>
<td>1.00629</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.85003</td>
<td>0.49480</td>
<td>1.00210</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td></td>
<td>-0.85000</td>
<td>0.49000</td>
<td>1.00000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

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

Fig. 4. The output estimation error $(\sigma_v^2 = 1.00^2, \delta_{ns} = 63.695\%)$

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 $\epsilon(k)$, we replace $v(k)$ and $\epsilon(k)$ with $\hat{v}(k)$ and $\hat{\epsilon}(k)$ 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.

4. CONCLUSIONS

A dual-rate recursive generalized extended forgetting gradient algorithm is used to estimate the parameters of the

10238
REFERENCES


