Personal answers to SI-EP-2016

BY WK.H

HITSZ

vvkc@foxmail.com

November 8, 2019

1.[7] Please give a definition of system identification, and then point out the four entities of system identification.

Answer 1.

(1) Identification can be defined as the determination of a mathematical model from the observed input and output data by minimizing some error criterion function.

(2) Four entities: data, the set of models, criterion, and optimization approaches.

....

2.[16] It is assumed that a, b, c, d and θ_i are unknown parameters, and $v(t)$ is the white noise. Please give the identification model of the following systems.

(1) $y(t) = \theta_1 u(t-1) + \theta_2 u^2(t-2) + \theta_3 u(t-1)u(t-2) + v(t)$

(2) $y(t) + \theta_1 y(t-1) = \theta_2 u(t-1) + \theta_3 u^2(t-2) + v(t)$

(3) $y(t) + \sin(\theta_1) y(t-1) = \theta_2 u(t-1) + \theta_3 \cos(t) + v(t)$

(4) $y(t) + y(t-1)y(t-2) = \theta_1 u(t-1) + \theta_2 u^2(t-2) + v(t)$

(5) $y(t) + ay(t-1) + by(t-2) = cu(t-1) + du(t-2) + v(t)$

(6) $y(t) = a + bt + ce^t + 4 + v(t)$

(7) $y(t) = ay(t-1) + by^2(t-2) + \frac{1}{c}[u(t) + du(t-1) + 2] + 4 + v(t)$

Answer 2.

(1) $y(t) = \theta_1 u(t-1) + \theta_2 u^2(t-2) + \theta_3 u(t-1)u(t-2) + v(t)$

$$\begin{cases} y(t) = \varphi^T(t)\theta + v(t) \\ \varphi(t) = [\ u(t-1) \ \ u^2(t-2) \ \ u(t-1)u(t-2)\ ]^T \\ \theta = [\ \theta_1 \ \ \theta_2 \ \ \theta_3\ ]^T \end{cases}$$

(2) $y(t) + \theta_1 y(t-1) = \theta_2 u(t-1) + \theta_3 u^2(t-2) + v(t)$

$$\begin{cases} y(t) = \varphi^T(t)\theta + v(t) \\ \varphi(t) = [\ -y(t-1) \ \ u(t-1) \ \ u^2(t-2)\ ]^T \\ \theta = [\ \theta_1 \ \ \theta_2 \ \ \theta_3\ ]^T \end{cases}$$

(3) $y(t) + \sin(\theta_1) y(t-1) = \theta_2 u(t-1) + \theta_3 \cos(t) + v(t)$

$$\begin{cases} y(t) = \varphi^T(t)\theta + v(t) \\ \varphi(t) = [\ -y(t-1) \ \ u(t-1) \ \ \cos(t)\ ]^T \\ \theta = [\ \sin(\theta_1) \ \ \theta_2 \ \ \theta_3\ ]^T \end{cases}$$

(4) $y(t)+y(t-1)y(t-2)=\theta_1 u(t-1)+\theta_2 u^2(t-2)+v(t)$

$$\begin{cases} y(t)+y(t-1)y(t-2)=\varphi^T(t)\theta+v(t) \\ \varphi(t)=[\ u(t-1)\ \ u^2(t-2)\ ]^T \\ \theta=[\ \theta_1\ \ \theta_2\ ]^T \end{cases}$$

(5) $y(t)+ay(t-1)+by(t-2)=cu(t-1)+du(t-2)+v(t)$

$$\begin{cases} y(t)=\varphi^T(t)\theta+v(t) \\ \varphi(t)=[\ -y(t-1)\ \ -y(t-2)\ \ u(t-1)\ \ u(t-2)\ ]^T \\ \theta=[\ a\ \ b\ \ c\ \ d\ ]^T \end{cases}$$

(6) $y(t)=a+bt+ce^t+4+v(t)$

$$\begin{cases} y(t)-4=\varphi^T(t)\theta+v(t) \\ \varphi(t)=[\ 1\ \ t\ \ e^t\ ]^T \\ \theta=[\ a\ \ b\ \ c\ ]^T \end{cases}$$

(7) $y(t)=ay(t-1)+by^2(t-2)+\frac{1}{c}[u(t)+du(t-1)+2]+4+v(t)$

$$\begin{cases} y(t)-4=\varphi^T(t)\theta+v(t) \\ \varphi(t)=[\ y(t-1)\ \ y^2(t-2)\ \ u(t)+2\ \ u(t-1)\ ]^T \\ \theta=\left[\ a\ \ b\ \ \frac{1}{c}\ \ \frac{d}{c}\ \right]^T \end{cases}$$

....

3.[16] Let $P^{-1}(t)=P^{-1}(t-1)+\varphi(t)\varphi^T(t)$, $\varphi^T(t)\varphi(t)\geq 0$, $\varphi(t)\in\mathbb{R}^n$. $P(0)=I$, I is the identity matrix with dimension n. Please prove the following results.

(1) $P(t)=P(t-1)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1+\varphi^T(t)P(t-1)\varphi(t)}$

(2) $P(t)\varphi(t)=\frac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)}$

(3) $P(t-1)\varphi(t)=\frac{P(t)\varphi(t)}{1-\varphi^T(t)P(t)\varphi(t)}$

Answer 3.

(1) By the matrix inversion lemma

$$(A+BC)^{-1}=A^{-1}-A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}$$

we can obtain

$$P(t)=P(t-1)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1+\varphi^T(t)P(t-1)\varphi(t)}$$

(2) From above, we can obtain

$$\begin{aligned} P(t)\varphi(t) &= P(t-1)\varphi(t)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)} \\ &= P(t-1)\varphi(t)\left(1-\frac{\varphi^T(t)P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)}\right) \\ &= \frac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)} \end{aligned}$$

(3) Similarly, by the matrix inversion lemma, we can obtain

$$P(t-1) = P(t) + \frac{P(t)\varphi(t)\varphi^T(t)P(t)}{1-\varphi^T(t)P(t)\varphi(t)}$$

Then, we can get

$$\begin{aligned} P(t-1)\varphi(t) &= P(t)\varphi(t) + \frac{P(t)\varphi(t)\varphi^T(t)P(t)\varphi(t)}{1-\varphi^T(t)P(t)\varphi(t)} \\ &= P(t)\varphi(t)\left(1+\frac{\varphi^T(t)P(t)\varphi(t)}{1-\varphi^T(t)P(t)\varphi(t)}\right) \\ &= \frac{P(t)\varphi(t)}{1-\varphi^T(t)P(t)\varphi(t)} \end{aligned}$$

....

4.[20] Given a third-order equation error model which is expressed by

$$y(t)+a_1y(t-1)+a_2y(t-2)+a_3y(t-3)=b_1u(t-1)+b_2u(t-2)+2\arctan(t)+v(t)$$

where $a_i, i=1,2,3$ and $b_j, j=1,2$ are the parameters to be identified. $\{y(t)\}$ and $\{u(t)\}$ are the output and input series of the system, and $\{v(t)\}$ is the white noise series with mean zero. Denote the parameter vector to be identified as $\theta = [\ a_1\ \ a_2\ \ a_3\ \ b_1\ \ b_2\ ]^T$.

(1) Please give the identification model of this system;

(2) Please give the expression of the least squares estimate $\hat{\theta}$ of θ when **batch** processing schedule is used (the data length is L);

(3) Please give the recursive least squares algorithm for this system;

(4) Please provide the expressions of the innovation and residual in the recursive least squares algorithm;

(5) Please establish a relation between the innovation and residual.

Answer 4.

(1)

$$\begin{cases} y(t)-2\arctan(t) = \varphi^T(t)\theta + v(t) \\ \varphi(t) = [\ y(t-1)\ \ y(t-2)\ \ y(t-3)\ \ u(t-1)\ \ u(t-2)\ ]^T \\ \theta = [\ a_1\ \ a_2\ \ a_3\ \ b_1\ \ b_2\ ]^T \end{cases}$$

(2)

$$\begin{cases} \hat{\theta}(L) = (H_L^TH_L)^{-1}H_L^TY_L \\ H_L = [\ \varphi(1)\ \ \varphi(2)\ \ \ldots\ \ \varphi(L)\ ]^T \\ Y_L = [\ y(1)-2\arctan(1)\ \ y(2)-2\arctan(2)\ \ \ldots\ \ y(L)-2\arctan(L)\ ]^T \end{cases}$$

(3)

$$\begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + G(t)[y(t)-\varphi^T(t)\hat{\theta}(t-1)-2\arctan(t)] \\ G(t) = \frac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)} \\ P(t) = [I - G(t)\varphi^T(t)]P(t-1), P(0) = p_0I \end{cases}$$

(4)

Innovation : $e(t) = y(t) - \varphi^T(t)\hat{\theta}(t-1) - 2\arctan(t)$;

Residual : $\varepsilon(t) = y(t) - \varphi^T(t)\hat{\theta}(t) - 2\arctan(t)$.

(5)

$$\begin{aligned}
\varepsilon(t) &= y(t) - \varphi^T(t)\hat{\theta}(t) - 2\arctan(t) \\
&= y(t) - \varphi^T(t)(\hat{\theta}(t-1) + G(t)e(t)) - 2\arctan(t) \\
&= y(t) - \varphi^T(t)\hat{\theta}(t-1) - 2\arctan(t) - \varphi^T(t)G(t)e(t) \\
&= e(t) - \varphi^T(t)G(t)e(t) \\
&= [1 - \varphi^T(t)P(t)\varphi(t)]e(t) \\
&= \left[1 - \frac{\varphi^T(t)P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)}\right]e(t) \\
&= \frac{1}{1+\varphi^T(t)P(t-1)\varphi(t)}e(t)
\end{aligned}$$

then we can conclude that $\varepsilon(t)$ and $e(t)$ have the same sign because $\varphi^T(t)P(t)\varphi(t) \geq 0$. Further, we can conclude that $\varphi^T(t)P(t)\varphi(t) \leq 1$.

....

5.[10] Given a third-order moving average model

$$y(t) = v(t) + d_1 v(t-1) + d_2 v(t-2) + d_3 v(t-3)$$

where $\{y(t)\}$ is the observation series, and $\{v(t)\}$ is the white noise series with mean zero. Let the identification model of this system be $y(t) = \varphi^T(t)\theta + v(t)$.

(1) Please give the expressions of vectors $\varphi(t)$ and θ;

(2) Please give the recursive extended least squares algorithm to estimate the parameter θ.

Answer 5.

(1) $\varphi(t) = [\ v(t-1)\ \ v(t-2)\ \ v(t-3)\ ]^T, \theta = [\ d_1\ \ d_2\ \ d_3\ ]^T$.

(2)

$$\begin{cases}
\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)] \\
L(t) = \frac{P(t-1)\hat{\varphi}(t)}{1+\hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)} \\
P(t) = [I - L(t)\hat{\varphi}^T(t)]P(t-1), P(0) = p_0 I \\
\hat{\varphi}(t) = [\ \hat{v}(t-1)\ \ \hat{v}(t-2)\ \ \hat{v}(t-3)\ ]^T \\
\hat{v}(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t)
\end{cases}$$

....

6.[11]

(1) Please explain the phenomenon of data saturation appearing in the recursive least squares identification algorithm;

(2) Please give a simple reason why such a phenomenon happens; (Hint: please refer to Problem 3)

(3) Please give a disadvantage resulted from data saturation.

(4) How to overcome data saturation? Please give a brief explain.

Answer 6.

(1) The new data has no contribution to improve the estimate $\hat{\theta}(t)$ of the parameter θ.

(2) Since $P(t) \geq 0$, by Problem 3, we can obtain that $P(t) \leq P(t-1)$. Further, we can get that $\lim_{t\to\infty} P(t) = 0$. It follows from this fact that the modified term $P(t)\varphi(t)[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)] \to 0$ when $t \to \infty$.

(3) If the system parameters to be identified are time-varying, then the estimation error can not be reduced as time goes by.

(4) One way is to introduce a factor for old data which makes the estimator more sensitive to recent samples. Another way is to use the most recent data with a certain data length only.

....

7.[20] Consider the following third-order output error autoregressive model

$$\begin{aligned}
y(t) &= \frac{B(z^{-1})}{A(z^{-1})}u(t) + \frac{D(z^{-1})}{C(z^{-1})}v(t) \\
A(z^{-1}) &= 1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} \\
B(z^{-1}) &= b_1 z^{-1} + b_2 z^{-2} \\
C(z^{-1}) &= 1 + c_1 z^{-1} + c_2 z^{-2} \\
D(z^{-1}) &= 1 + d_1 z^{-1} + d_2 z^{-2}
\end{aligned}$$

where $a_i, i = 1, 2, 3$ and $b_j, c_j, d_j, j = 1, 2$ are the parameters to be identified; $\{y(t)\}$ and $\{u(t)\}$ are the output and input series of the system, and $\{v(t)\}$ is the white noise series with mean zero. In addition, let the system model be

$$x(t) = \frac{B(z^{-1})}{A(z^{-1})}u(t)$$

and the noise model be

$$w(t) = \frac{D(z^{-1})}{C(z^{-1})}v(t)$$

(1) Please give the identification model of the considered output error autoregressive model according to outputs of the system model and the noise model;

(2) Please give the auxiliary model based recursive generalized least squares algorithm to estimate the parameters $a_i, i = 1, 2, 3$ and $b_j, c_j, d_j, j = 1, 2$. The detailed derivation should be provided.

Answer 7.

(1) $y(t) = \varphi^T(t)\theta + v(t)$ where

$$\begin{cases}
\varphi(t) = \begin{bmatrix} \varphi_s(t) \\ \varphi_n(t) \end{bmatrix}, \theta = \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix} \\
\varphi_s(t) = [\ -x(t-1) \quad -x(t-2) \quad -x(t-3) \quad u(t-1) \quad u(t-2)\ ]^T \\
\varphi_n(t) = [\ -\omega(t-1) \quad -\omega(t-2) \quad v(t-1) \quad v(t-2)\ ]^T \\
\theta_s = [\ a_1 \quad a_2 \quad a_3 \quad b_1 \quad b_2\ ]^T \\
\theta_n = [\ c_1 \quad c_2 \quad d_1 \quad d_2\ ]^T
\end{cases}$$

(2)

In order to estimate unknown terms in $\varphi(t)$, the following relations are used

$$\begin{aligned}
v(t) &= y(t) - \varphi^T(t)\theta \\
x(t) &= \varphi_s^T(t)\theta_s \\
\omega(t) &= \varphi_n^T(t)\theta_n + v(t)
\end{aligned}$$

By replacing the unknown terms with their estimations, we can obtain

$$y(t) = \hat{\varphi}^T(t)\hat{\theta}(t) + v(t)$$

where

$$\begin{cases} \hat{\varphi}(t) = \left[\begin{array}{c} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{array} \right], \hat{\theta}(t) = \left[\begin{array}{c} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{array} \right] \\ \hat{\varphi}_s(t) = [\ -x_a(t-1) \quad -x_a(t-2) \quad -x_a(t-3) \quad u(t-1) \quad u(t-2)\ ]^T \\ \hat{\varphi}_n(t) = [\ -\hat{\omega}(t-1) \quad -\hat{\omega}(t-2) \quad \hat{v}(t-1) \quad \hat{v}(t-2)\ ]^T \\ \hat{v}(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t) \\ x_a(t) = \hat{\varphi}_s^T(t)\hat{\theta}_s(t) \\ \hat{\omega}(t) = y(t) - x_a(t) \end{cases} \qquad (*)$$

Combing the definition in $(*)$, the auxiliary model based recursive generalized least squares algorithm can be given as follows

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-1) + L(t)\big[\, y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)\big] \\ L(t) &= \frac{P(t-1)\hat{\varphi}(t)}{1 + \hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)} \\ P(t) &= [I - L(t)\hat{\varphi}^T(t)]P(t-1), P(0) = p_0 I \end{aligned}$$