## 4 System identification; Stochastic identifier

### 4.1 Least square estimation

We considet following system;

$$
\begin{equation*}
y(k)=\frac{B\left(q^{-1}\right)}{A\left(q^{-1}\right)} u(k)+n(k) \tag{1}
\end{equation*}
$$

Here, we assume $\mathrm{m}=\mathrm{n}$ (dimension of A and B are equal).
Then, $\mathrm{y}(\mathrm{k})$ is represented by following equations.

$$
\begin{gather*}
y(k)=-a_{1} y(k-1)-a_{2} y(k-2)-\cdots-a_{n} y(k-n)+b_{0} u(k)+b_{1} u(k-1)+\cdots+b_{n} u(k-n)+w(k)  \tag{3}\\
w(k)=A\left(q^{-1}\right) n(k)=n(k)+a_{1} n(k-1)+\cdots+a_{n} n(k-n) \tag{2}
\end{gather*}
$$

Our aim is estimating system parameters $a_{i}$ and $b_{i}$ by observed $u(k)$ and $y(k)$ values $(k=1,2, \cdots, N)$.

At first, we define $\theta$ and $z(k)$.

$$
\begin{align*}
\theta^{T} & \left.=\left[\begin{array}{cccccccc}
-a_{1} & -a_{2} & \cdots & -a_{n} & b_{0} & b_{1} & \cdots & b_{n} \\
z^{T}(k) & =\left[\begin{array}{cccc}
y(k-1) & y(k-2) & \cdots & y(k-n) \\
u(k) & u(k-1) & \cdots & u(k-n)
\end{array}\right]
\end{array}\right] \begin{array}{ll} 
& (k-1
\end{array}\right) \tag{4}
\end{align*}
$$

Then, equation(2) becomes

$$
\begin{equation*}
y(k)=z^{T}(k) \theta+w(k) \tag{5}
\end{equation*}
$$

We also define vectors $y, w$ and $Z$ as follows.

Then, output $y$ is derived by (7).

$$
\begin{equation*}
y=Z \theta+w \tag{7}
\end{equation*}
$$

Because $w(k)$ represents noise, so we want to minimize $J$.

$$
\begin{equation*}
J=\sum_{k=1}^{N} w^{2}(k)=w^{T} w=(y-Z \theta)^{T}(y-Z \theta) \tag{8}
\end{equation*}
$$

$\frac{\partial J}{\partial \theta}$ is

$$
\begin{equation*}
\frac{\partial J}{\partial \theta}=-2 Z^{T}(y-Z \theta) \tag{9}
\end{equation*}
$$

For $\frac{\partial J}{\partial \theta}=0$, estimated $\theta(\hat{\theta})$ is

$$
\begin{align*}
\hat{\theta} & =\left(Z^{T} Z\right)^{-1} Z^{T} y  \tag{10}\\
& =\left[\sum_{k=1}^{N} z(k) z^{T}(k)\right]^{-1} \sum_{k=1}^{N} z(k) y(k) .
\end{align*}
$$

### 4.1.1 Consider estimated $\theta$

Let us consider estimation error. When we define $\epsilon_{\theta}=\hat{\theta}-\theta$,

$$
\begin{equation*}
\epsilon_{\theta}=\left(Z^{T} Z\right)^{-1} Z^{T} w \tag{11}
\end{equation*}
$$

If $w(k)$ is a white noise,

$$
\begin{equation*}
E[w(k)]=0, E\left[w(i) w^{T}(j)\right]=Q \delta_{i j} \tag{12}
\end{equation*}
$$

So, if $Z$ is independent from $w, E\left[\epsilon_{\theta}\right]=0$.

### 4.1.2 Online identification

We have already observed $N$ data. Now, we obtain $N+1$ th data. Let us assume $y_{N+1}, Z_{N+1}, P(N)$ and $q(N)$ as follows.

$$
\begin{gather*}
y_{N+1}=\left[\begin{array}{c}
y_{N} \\
\cdots \\
y(N+1)
\end{array}\right], Z_{N+1}=\left[\begin{array}{c}
Z_{N} \\
\cdots \\
z^{T}(N+1)
\end{array}\right]  \tag{13}\\
P(N)=\left[Z_{N}^{T} Z_{N}\right]^{-1}, q(N)=Z_{N}^{T} y_{N} \tag{14}
\end{gather*}
$$

Then,

$$
\begin{gather*}
P^{-1}(N+1)=P^{-1}(N)+z(N+1) z^{T}(N+1)  \tag{15}\\
q(N+1)=q(N)+z(N+1) y(N+1) \tag{16}
\end{gather*}
$$

By the matrix inversion lemma,

$$
\begin{equation*}
P(N+1)=P(N)-P(N) \frac{z(N+1) z^{T}(N+1)}{1+z^{T}(N+1) P(N) z(N+1)} P(N) \tag{17}
\end{equation*}
$$

Then, by equation (10),

$$
\begin{align*}
\hat{\theta}(N+1)= & P(N+1) q(N+1)=  \tag{18}\\
& {\left[P(N)-P(N) \frac{z(N+1) z^{T}(N+1)}{1+z^{T}(N+1) P(N) z(N+1)} P(N)\right] } \\
& \times\left[Z_{N}^{T} y_{N}+z(N+1) y(N+1)\right] \tag{19}
\end{align*}
$$

Therefore, we can calculate parameters successively. $(N \rightarrow k)$

$$
\begin{align*}
\hat{\theta}(k+1) & =\hat{\theta}(k)+K(k+1)\left[y(k+1)-z^{T}(k+1) \hat{\theta}(k)\right]  \tag{20}\\
K(k+1) & =\frac{P(k) z(k+1)}{1+z^{T}(k+1) P(k) z(k+1)} \\
P(k+1) & =\left[I-K(k+1) z^{T}(k+1)\right] P(k)
\end{align*}
$$

Additionally, by $K(N+1)=P(N+1) z(N+1)$,

$$
\begin{align*}
\hat{\theta}(k+1) & =\hat{\theta}(k)+P(k+1) z(k+1)\left[y(k+1)-z^{T}(k+1) \hat{\theta}(k)\right]  \tag{21}\\
P(k+1) & =P(k)-P(k) \frac{z(k+1) z^{T}(k+1)}{1+z^{T}(k+1) P(k) z(k+1)} P(k)
\end{align*}
$$

### 4.2 Weighted least square estimation

We can add weight matrix $Q$ to the evaluation function $J$.

$$
\begin{equation*}
J=w^{T} Q w=(y-Z \theta)^{T} Q(y-Z \theta) \tag{22}
\end{equation*}
$$

In this case, estimated $\theta(\hat{\theta})$ becomes

$$
\begin{equation*}
\hat{\theta}=\left(Z^{T} Q Z\right)^{-1} Z^{T} Q y \tag{23}
\end{equation*}
$$

### 4.3 Extended least square method

When $w(k)$ is not the white noise, former methods cannot cancel the bias error.Here, we introduce extended least square method to estimate not only $\theta$ but also noise properties.

We considet following system;

$$
\begin{equation*}
y(k)=\frac{B\left(q^{-1}\right)}{A\left(q^{-1}\right)} u(k)+n(k) \tag{24}
\end{equation*}
$$

Then,

$$
\begin{align*}
A\left(q^{-1}\right) y(k) & =B\left(q^{-1}\right) u(k)+w(k)  \tag{25}\\
w(k) & =A\left(q^{-1}\right) n(k)
\end{align*}
$$

We re-define $w(k)$ by equation $(26)$, where $m(k)$ indicates the white noise, and $C\left(q^{-1}\right)$ represents characteristics of the noise.

$$
\begin{gather*}
w(k)=\frac{1}{C\left(q^{-1}\right)} m(k)  \tag{26}\\
C\left(q^{-1}\right)=1+\sum_{i=1}^{p} c_{i} q^{-i} \tag{27}
\end{gather*}
$$

Then,

$$
\begin{gather*}
w(k)+\sum_{i=1}^{p} c_{i} w(k-i)=m(k)  \tag{28}\\
y(k)=-\sum_{i=1}^{n} a_{i} y(k-i)+\sum_{i=0}^{m} b_{i} u(k-i)-\sum_{i=1}^{p} c_{i} w(k-i)+m(k) \tag{29}
\end{gather*}
$$

Here, we define vectors $\theta, z_{m}(k), c, w(k)$ as follows.

Then,

$$
\begin{equation*}
y(k)=z_{m}^{T}(k) \theta+w^{T}(k) c+m(k) \tag{31}
\end{equation*}
$$

Additionally, we define $Z, W, m, y$ that combiles observed data $k=1,2, \cdots, N$.

$$
\begin{align*}
Z^{T} & =\left[\begin{array}{cccc}
z_{m}(1) & z_{m}(2) & \cdots & z_{m}(N)
\end{array}\right] \\
W^{T} & =\left[\begin{array}{cccc}
w(1) & w(2) & \cdots & w(N)
\end{array}\right]  \tag{32}\\
m^{T} & =\left[\begin{array}{cccc}
m(1) & m(2) & \cdots & m(N)
\end{array}\right] \\
y^{T} & =\left[\begin{array}{cccc}
y(1) & y(2) & \cdots & y(N)
\end{array}\right]
\end{align*}
$$

Then,

$$
\begin{equation*}
y=Z \theta+W c+m \tag{33}
\end{equation*}
$$

Finally, we transform equation (33) into

$$
\begin{equation*}
m=y-\Omega \phi \tag{34}
\end{equation*}
$$

where

$$
\Omega=[Z \vdots W], \phi=\left[\begin{array}{c}
\theta  \tag{35}\\
\cdots \\
c
\end{array}\right] .
$$

We introduce $\Omega$, called extended matrix, so that we can estimate both $\theta$ and $c$ simultaneously.

$$
\begin{gather*}
\hat{\phi}=\left(\Omega^{T} \Omega\right)^{-1} \Omega^{T} y  \tag{36}\\
\hat{\theta}(k)=\left(Z^{T}(k) Z(k)\right)^{-1}\left[Z^{T}(k) y(k)-Z^{T}(k) W(k) \hat{c}(k)\right]  \tag{37}\\
\hat{c}(k)=\left(W^{T}(k) W(k)\right)^{-1} W^{T}(k)[y(k)-Z(k) \hat{\theta}(k)] \tag{38}
\end{gather*}
$$

### 4.3.1 On-line estimation

Equation (31) gives $w(k)=y(k)-z_{m}^{T}(k) \hat{\theta}(k)$. Then,

$$
\begin{align*}
\hat{\phi}(k+1) & =\hat{\phi}(k)+K(k+1)\left[y(k+1)-\Omega^{T}(k+1) \hat{\phi}(k)\right]  \tag{39}\\
K(k+1) & =\frac{P(k) \Omega^{T}(k+1)}{I+\Omega(k+1) P(k) \Omega(k+1)} \\
P(k+1) & =\left[I-K(k+1) \Omega^{T}(k+1)\right] P(k)
\end{align*}
$$

