Department of Engineering Cybernetics

TTK4215 System Identification and Adaptive Control
Solution 2
Prediction-Error Methods – The Least-Squares Method

Problem 1 (Optimal Predictions)

a) The prediction error is

$$\varepsilon(t) = -ay(t-1) + bu(t-1) + e(t) + ce(t-1) - \hat{y}(t). \tag{1}$$

The variance of the prediction error is

$$\begin{aligned} Var(\varepsilon) &= E\varepsilon^2(t) - [E\varepsilon(t)]^2 \\ &= E\left(-ay(t-1) + bu(t-1) + e(t) + ce(t-1) - \hat{y}(t)\right)^2 - [E\varepsilon(t)]^2 \\ &= E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)^2 \\ &\quad +2E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)e(t) \\ &\quad +Ee^2(t) - [E\varepsilon(t)]^2. \end{aligned} \tag{2}$$

Since $e(t)$ is white noise, it is independent of all the other signals, so that

$$E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)e(t) = 0. \tag{3}$$

We get

$$\begin{aligned} Var(\varepsilon) &= E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)^2 + \lambda - [E\varepsilon(t)]^2 \\ &= E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)^2 + \lambda \\ &\quad -[E\varepsilon(t) - Ee(t)]^2 \\ &= E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)^2 + \lambda \\ &\quad -[E\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right)]^2 \\ &= Var\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right) + \lambda. \end{aligned} \tag{4}$$

Since all variances are positive, that is $Var\left(-ay(t-1) + bu(t-1) + ce(t-1) - \hat{y}(t)\right) \geq 0$, it follows that

$$Var(\varepsilon) \geq \lambda. \tag{5}$$

b) From (4) and (5) it is clear that

$$\hat{y}(t) = -ay(t-1) + bu(t-1) + ce(t-1) \tag{6}$$

constitutes an optimal predictor, since it gives equality in (5). However, it is not implementable, due to the term involving $e(t-1)$. However, if we use (34) from the problem set, we may write

$$e(t-1) = y(t-1) + ay(t-2) - bu(t-2) - ce(t-2), \tag{7}$$

which substituted into (6) gives

$$\begin{aligned}\hat{y}(t) &= -ay(t-1)+bu(t-1)+c(y(t-1)+ay(t-2)-bu(t-2)-ce(t-2)) \\ &= -c[-ay(t-2)+bu(t-2)+ce(t-2)]+(c-a)y(t-1)+bu(t-1).\end{aligned}$$

Noticing that the expression within the brackets in the above equation is $\hat{y}(t-1)$, we obtain

$$\hat{y}(t) = -c\hat{y}(t-1)+(c-a)y(t-1)+bu(t-1),$$

as an alternative way of expressing the optimal predictor (6).

c) Applying the backward shift operator, we have that

$$\hat{y}(t) = -cq^{-1}\hat{y}(t)+bq^{-1}u(t)+(c-a)q^{-1}y(t). \tag{8}$$

Thus, we get

$$\left(1+cq^{-1}\right)\hat{y}(t) = bq^{-1}u(t)+(c-a)q^{-1}y(t). \tag{9}$$

so

$$\hat{y}(t) = \frac{bq^{-1}}{1+cq^{-1}}u(t)+\frac{(c-a)q^{-1}}{1+cq^{-1}}y(t). \tag{10}$$

d)

$$\hat{y}(t) = y(t)-e(t) = G(q)u(t)+H(q)e(t)-e(t). \tag{11}$$

Since the noise model is invertible, we can write

$$e(t) = H^{-1}(q)(y(t)-G(q)u(t)). \tag{12}$$

Substituting (12) into (11) gives

$$\begin{aligned}\hat{y}(t) &= G(q)u(t)+(H(q)-I)H^{-1}(q)(y(t)-G(q)u(t)) \\ &= G(q)u(t)+\left(I-H^{-1}(q)\right)(y(t)-G(q)u(t)) \\ &= G(q)u(t)+\left(I-H^{-1}(q)\right)y(t)-G(q)u(t)+H^{-1}(q)G(q)u(t) \\ &= H^{-1}(q)G(q)u(t)+\left(I-H^{-1}(q)\right)y(t).\end{aligned} \tag{13}$$

Problem 2 (Least-Squares Method)

a) The problem is quadratic in θ, so the global optimum is found easily found by differentiating with respect to θ to obtain

$$\begin{aligned}\frac{\partial}{\partial\theta}\frac{1}{2N}\sum_{t=1}^{N}\left(y(t)-\varphi^T(t)\theta\right)^2 &= \frac{1}{N}\sum_{t=1}^{N}-\varphi(t)\left(y(t)-\varphi^T(t)\theta\right) \\ &= -\frac{1}{N}\sum_{t=1}^{N}\left(\varphi(t)y(t)-\varphi(t)\varphi^T(t)\theta\right).\end{aligned} \tag{14}$$

and setting the result equal to zero

$$-\frac{1}{N}\sum_{t=1}^{N}\left(\varphi(t)y(t)-\varphi(t)\varphi^T(t)\theta\right) = 0.$$

This gives

$$\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,\varphi^{T}(t)\,\theta=\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,y(t)\,. \tag{15}$$

Provided

$$\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,\varphi^{T}(t) \tag{16}$$

is invertible, we get

$$\hat{\theta}_{N}^{LS}=\left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,\varphi^{T}(t)\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,y(t)\,. \tag{17}$$

The regression vector $\varphi(t)$ needs to be sufficiently varied as a function of time to make the matrix (16) invertible. Typically, this is obtained by applying a sufficiently rich input signal $u(t)$.

b) We have

$$\left(1+a_1q^{-1}+a_2q^{-2}\right)\hat{y}(t)=\left(b_1q^{-1}+b_2q^{-2}+b_3q^{-3}\right)u(t)+v(t)\,, \tag{18}$$

so

$$y(t)=-a_1y(t-1)-a_2y(t-2)+b_1u(t-1)+b_2u(t-2)+b_3u(t-3)+v(t)\,. \tag{19}$$

So, defining

$$\varphi(t) = \begin{bmatrix} -y(t-1) & -y(t-2) & u(t-1) & u(t-2) & u(t-3) \end{bmatrix}^{T}, \tag{20}$$

$$\theta = \begin{bmatrix} a_1 & a_2 & b_1 & b_2 & b_3 \end{bmatrix}^{T}, \tag{21}$$

provides the desired form $\hat{y}(t)=\varphi^{T}(t)\,\theta$. The following MATLAB code generates the data, and solves this assignment (you were not supposed to generate data, of course).

```
% Set true parameters
a=[0.5;0.2];
b=[0.3;-0.2;0.1];

% Generate noise, and input signal.
e=random('Normal',0,1,10000,1);
U=random('Normal',0,1,10000,1);

% Set previous data (u and y).
prevu=[0;0;0];
prevy=[0;0];
N=length(e);
Y=zeros(N,1);
```

```
% Generate output.
for t=1:N,
% Compute current y
   y=-a'*prevy+b'*prevu+e(t);
% Store it.
   Y(t)=y;
% Set previous data:
   prevu=[U(t);prevu(1:2)];
   prevy=[y;prevy(1)];
end

% Store data
save Data_Assignment2 U Y;

%Compute LS-estimate from U and Y.

% Initialize the sums of the LS equation
SUM1=zeros(5,5);
SUM2=zeros(5,1);

% Initilize storage
THETA=[];

% Time loop.
for t=1:N,
    % Regressor
    if (t==1)
       phi=[0;0;0;0;0];
    elseif(t==2)
       phi=[-Y(1);0;U(1);0;0];
    elseif(t==3)
       phi=[-Y(2);-Y(1);U(2);U(1);0];
    else
       phi=[-Y(t-1);-Y(t-2);U(t-1);U(t-2);U(t-3)];
    end

    % Compute sums
    SUM1=SUM1+phi*phi';
    SUM2=SUM2+phi*Y(t);

    % If j is larger or equal to 10, compute and store estimate.
    if (t>=10)
       theta=inv(SUM1/t)*SUM2/t;
       THETA=[THETA,theta];
    end
end
```

```
% Plot estimate as function of N.
plot(THETA');

% Display the final estimate
theta
```

c) In the multivariable case we have

$$V_N\left(\theta,\mathbb{Z}^N\right)=\frac{1}{N}\sum_{t=1}^{N}\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right)^T\Lambda^{-1}\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right)/2. \tag{22}$$

As above, we differentiate with respect to θ, and set the result equal to 0. We get

$$\begin{aligned}\frac{\partial}{\partial\theta}V_N &= \frac{1}{N}\sum_{t=1}^{N}-\frac{1}{2}\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right)^T\Lambda^{-1}\varphi^T\left(t\right)-\left(\frac{1}{2}\varphi\left(t\right)\Lambda^{-1}\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right)\right)^T\\ &= -\frac{1}{N}\sum_{t=1}^{N}\left(\varphi\left(t\right)\Lambda^{-1}\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right)\right)^T, \end{aligned}\tag{23}$$

so

$$\frac{1}{N}\sum_{t=1}^{N}\left(\varphi\left(t\right)\Lambda^{-1}\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right)\right)=0 \tag{24}$$

which gives

$$\left(\frac{1}{N}\sum_{t=1}^{N}\varphi\left(t\right)\Lambda^{-1}\varphi^T\left(t\right)\right)\theta=\frac{1}{N}\sum_{t=1}^{N}\varphi\left(t\right)\Lambda^{-1}y\left(t\right). \tag{25}$$

Provided the matrix in front θ above is invertible, we obtain

$$\hat{\theta}_{LS}^{N}=\left(\frac{1}{N}\sum_{t=1}^{N}\varphi\left(t\right)\Lambda^{-1}\varphi^T\left(t\right)\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi\left(t\right)\Lambda^{-1}y\left(t\right). \tag{26}$$

d) In the case of measurement weighting, we get

$$\frac{\partial}{\partial\theta}V_N=\sum_{t=1}^{N}2\beta\left(N,t\right)\varphi\left(t\right)\left(y\left(t\right)-\varphi^T\left(t\right)\theta\right), \tag{27}$$

so we get

$$\sum_{t=1}^{N}\beta\left(N,t\right)\varphi\left(t\right)\varphi^T\left(t\right)\theta=\sum_{t=1}^{N}\beta\left(N,t\right)\varphi\left(t\right)y\left(t\right). \tag{28}$$

Thus, provided the matrix in front of θ is invertible, we get

$$\hat{\theta}_{N}^{WLS}=\left(\sum_{t=1}^{N}\beta\left(N,t\right)\varphi\left(t\right)\varphi^T\left(t\right)\right)^{-1}\sum_{t=1}^{N}\beta\left(N,t\right)\varphi\left(t\right)y\left(t\right). \tag{29}$$

e) We have

$$\begin{aligned}
\tilde{\theta}_N &= \hat{\theta}_N^{LS} - \theta^* \\
&= \left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,y(t) - \theta^* \\
&= \left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\left(\varphi^T(t)\,\theta^* + v_0(t)\right) - \theta^* \\
&= \left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right)\theta^* - \theta^* \\
&\quad + \left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,v_0(t) \\
&= \left(\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\,v_0(t)\,. \qquad (30)
\end{aligned}$$

f) We have

$$R^* = \bar{E}\varphi(t)\,\varphi^T(t) = \bar{E}\begin{bmatrix} y^2(t-1) & -y(t-1)\,u(t-1) \\ -y(t-1)\,u(t-1) & u^2(t-1) \end{bmatrix}. \qquad (31)$$

Since $y(t-1)$ is independent of $u(t-1)$, the off-diagonal terms in (31) are zero. Also, $\bar{E}u^2(t-1) = \bar{E}u^2(t) = 1$. For the last term, we have

$$\bar{E}y^2(t-1) = \bar{E}y^2(t) \qquad (32)$$

since $y(t)$ is stationary since u and v are. We have

$$\begin{aligned}
\bar{E}y^2(t) &= \bar{E}\left(-a_0 y(t-1) + b_0 u(t-1) + e_0(t) + c_0 e(t-1)\right)^2 && (33)\\
&= a_0^2\bar{E}y^2(t-1) + b_0^2\bar{E}u^2(t-1) + \bar{E}e_0^2(t) + c_0^2\bar{E}e_0^2(t-1) && (34)\\
&\quad -2a_0b_0\bar{E}y(t-1)\,u(t-1) - 2a_0\bar{E}y(t-1)\,e_0(t) && (35)\\
&\quad -2a_0c_0\bar{E}y(t-1)\,e_0(t-1) + 2b_0\bar{E}u(t-1)\,e_0(t) && (36)\\
&\quad +2b_0c_0\bar{E}u(t-1)\,e_0(t-1) + 2c_0\bar{E}e_0(t)\,e_0(t-1)\,. && (37)
\end{aligned}$$

Due to the fact that $u(t-1)$ and $e_0(t)$ are independent of the other signals, we obtain

$$\begin{aligned}
\bar{E}y^2(t) &= a_0^2\bar{E}y^2(t-1) + b_0^2 + 1 + c_0^2 && (38)\\
&\quad -2a_0c_0\bar{E}y(t-1)\,e_0(t-1)\,. && (39)
\end{aligned}$$

Due to (32) and

$$\bar{E}y(t-1)\,e_0(t-1) = \bar{E}\left(-a_0 y(t-2) + b_0 u(t-2) + e_0(t-1) + c_0 e_0(t-2)\right)e_0(t-1) = 1, \qquad (40)$$

we get

$$\bar{E}y^2(t) = a_0^2\bar{E}y^2(t) + b_0^2 + 1 + c_0^2 - 2a_0c_0. \tag{41}$$

Solving for $\bar{E}y^2(t)$, we get

$$\bar{E}y^2(t) = \frac{b_0^2 + c_0^2 - 2a_0c_0 + 1}{1 - a_0^2}. \tag{42}$$

So, we get

$$(R^*)^{-1} = \begin{bmatrix} \frac{1-a_0^2}{b_0^2+c_0^2-2a_0c_0+1} & 0 \\ 0 & 1 \end{bmatrix}. \tag{43}$$

For f^*, we get

$$\begin{aligned} f^* &= \bar{E}\begin{bmatrix} -y(t-1)(e_0(t) + c_0e_0(t-1)) \\ u(t-1)(e_0(t) + c_0e_0(t-1)) \end{bmatrix} \\ &= \bar{E}\begin{bmatrix} -c_0y(t-1)e_0(t-1) \\ 0 \end{bmatrix} = \begin{bmatrix} -c_0 \\ 0 \end{bmatrix}. \end{aligned} \tag{44}$$

Finally, we obtain

$$\lim_{N\to\infty}\tilde{\theta}_N = \begin{bmatrix} \frac{1-a_0^2}{b_0^2+c_0^2-2a_0c_0+1} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -c_0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{\left(1-a_0^2\right)c_0}{b_0^2+c_0^2-2a_0c_0+1} \\ 0 \end{bmatrix}. \tag{45}$$