7 Observer / *Kalman* filter identification; OKID:

7.1 Introduction:

- What to ID: state space model.
- Why to ID:

1. For modal parameters of a large flexible structure; damping, frequencies, mode shape and mode amplitude.
2. Controller design with state feedback.

- How to ID:

$$\left.\begin{array}{l} FRF \xrightarrow{IDFT} \\ {}^{I}\!/\!{}_{O}\ history \xrightarrow[\text{Time domain algorithm}]{\text{FIR model; RLS, lattice filter}} \end{array}\right\} \begin{array}{l} Markov \\ parameters \end{array} \xrightarrow{ERA} \begin{array}{l} state \\ space \end{array}.$$

- What is the problem for OKID to solve?

The parameters length of FIR model is too long to solve with least square approach.

- How does OKID solve the above problem?

Identify an asymptotically stable observer to form a stable state-space model of observer and then find the state space model of system from observer model.

- When system $\begin{bmatrix} A & B & C & D \end{bmatrix}$ is asymptotic stable, its state space is represented in impulse response from $\mathcal{Y}_{m\times\ell} = \bar{Y}_{m\times r\ell}\left(\mathcal{U}_p\right)_{r\ell\times\ell}$ where $\mathcal{Y} = \begin{bmatrix} y(0) & y(1) & \cdots & y(\ell-1) \end{bmatrix}$, $\bar{Y} = \begin{bmatrix} D & CB & \cdots & CA^{\ell-2}B \end{bmatrix}$ and

$$\mathcal{U}_p = \begin{bmatrix} u(0) & u(1) & u(2) & \cdots & u(\ell-1) \\ 0 & u(0) & u(1) & \cdots & u(\ell-2) \\ \vdots & \ddots & u(0) & \cdots & u(\ell-3) \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & u(0) \end{bmatrix}.$$ For large ℓ, $\Rightarrow CA^{\ell-2}B \to 0$. It's larger for light damping and smaller for heavy damping.

- State space observer for $\begin{bmatrix} A & B & C & D \end{bmatrix}$ system.

$\bar{x}(k+1) = A\bar{x}(k) + Bu(k) - G\left[y(k) - \bar{y}(k)\right]$

$\bar{y}(k) = C\bar{x}(k) + Du(k)$

$\Rightarrow \bar{x}(k+1) = (A+GC)\bar{x}(k) + (B+GD)u(k) - Gy(k)$

$= (A+GC)\bar{x}(k) + \begin{bmatrix} B+GD & -G \end{bmatrix} \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \triangleq \bar{A}\bar{x}(k) + \bar{B}\bar{u}(k).$

- Impulse response model: $\mathcal{Y}_{m\times\ell} = \bar{Y}_{m\times[(m+r)(\ell-1)+r]} \left(\bar{\mathcal{U}}_p\right)_{[(m+r)(\ell-1)+r]\times\ell}$,

where $\mathcal{Y} = \left[y(0) \quad y(1) \quad \cdots \quad y(p) \quad \cdots \quad y(\ell-1)\right]$,

$\bar{Y} = \begin{bmatrix} D & C\bar{B} & C\bar{A}\bar{B} & \cdots & C\bar{A}^{p-2}\bar{B} & \cdots & C\bar{A}^{\ell-2}\bar{B} \end{bmatrix}$; observer *Markov* parameters and

$$\bar{\mathcal{U}}_p = \begin{bmatrix} u(0) & u(1) & u(2) & \cdots & u(p) & \cdots & u(\ell-1) \\ 0 & \bar{u}(0) & \bar{u}(1) & \cdots & \bar{u}(p-1) & \cdots & \bar{u}(\ell-2) \\ \vdots & \ddots & \bar{u}(0) & \cdots & \bar{u}(p-2) & \cdots & \bar{u}(\ell-3) \\ \vdots & & \ddots & \ddots & \vdots & & \vdots \\ \vdots & & & \ddots & \bar{u}(0) & \cdots & \bar{u}(\ell-p-1) \\ \vdots & & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \bar{u}(0) \end{bmatrix}.$$

Again, if we need the term ℓ for $CA^{\ell-2}B \to 0$ to be small then $\bar{A}$ need to be highly damped. $\Rightarrow$ Eigenvalues be small. It means that we choose G for $\bar{A} = A + GC$ to make $\bar{A}$ having high damping.

- Objective of observer: $\bar{x}(k) \to x(k)$

$\because \bar{x}(k+1) = A\bar{x}(k) + Bu(k) - G\left[y(k) - \bar{y}(k)\right] = A\bar{x}(k) + Bu(k)$

$-GCx(k) + GC\bar{x}(k) - GDu(k) + GDu(k)$ and $x(k+1) = Ax(k) + Bu(k)$

$\therefore x(k+1) - \bar{x}(k+1) = A\left[x(k) - \bar{x}(k)\right] + GC\left[x(k) - \bar{x}(k)\right]$

$= \bar{A}\left[x(k) - \bar{x}(k)\right]$. $\bar{A}$ having highly damping is helpful to the perfor–mance of observer. The objective of observer requires same condition for $\bar{A}$ to be in order to shorten length of $\bar{Y}$.

- Dead beat observer: If G makes the eigenvalues of $\bar{A} = A + GC$ to be all zero. $\Rightarrow \bar{A}^n = 0$, where n is the order of $\bar{A}$. $\Rightarrow \bar{x}(n) = x(n)$, $\forall x(0) \neq \bar{x}(0)$. It means that the state estimation error of a dead beat observer will converge on zeros after n steps.

- Initial condition effect: Output response is affected by initial state and input; $\bar{\mathbf{y}} = C\bar{A}^p X + \bar{Y}\bar{\bar{\mathbf{u}}}_p$, where $\mathbf{y} = \left[y(p) \quad y(p+1) \quad \cdots \quad y(\ell-1)\right]$, $X = \left[x(0) \quad x(1) \quad \cdots \quad x(\ell-p-2)\right]$, $\bar{Y} = \left[D \quad C\bar{B} \quad \cdots \quad C\bar{A}^{p-1}\bar{B}\right]$ and

$$\bar{\bar{\mathbf{u}}}_p = \begin{bmatrix} u(p) & u(p+1) & \cdots & u(\ell-1) \\ \bar{u}(p-1) & \bar{u}(p) & \cdots & \bar{u}(\ell-2) \\ \vdots & \vdots & & \vdots \\ \bar{u}(0) & \bar{u}(1) & \cdots & \bar{u}(\ell-p-1) \end{bmatrix}$$

. If $\bar{A}^p \to 0$ $\Rightarrow C\bar{A}^p X \to 0$ $\Rightarrow \mathbf{y} \cong \bar{Y}\bar{\bar{\mathbf{u}}}_p$ then we will omit the initial condition as $\bar{A}$ having highly damping.

- OKID from the history of $\bar{u}(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \xrightarrow[RLS]{LS}$ *Markov* parameters $\bar{Y}\ \cancel{\xrightarrow{ERA}}\ \left[\bar{A} \quad \bar{B} \quad C \quad D\right]$. It from the history of $\bar{u}(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \xrightarrow{LS}$ observer *Markov* parameters $\bar{Y} \xrightarrow[\bar{B}=\left[B+GD \quad -G\right]]{\bar{A}=A+GC}$

$$\begin{cases} Y : \text{system Markov parameters} \xrightarrow{ERA} \begin{bmatrix} A & B & C \end{bmatrix} \\ Y^0 : \begin{matrix} \text{observer gain} \\ \text{markov parameters} \end{matrix} \xrightarrow{ERA} G \end{cases}.$$

7.2 System *Markov* parameters:

- Compute system *Markov* parameters from given observer *Markov* parameters:

$\because \bar{Y}_0 = D,\ \bar{Y}_k = C\bar{A}^{k-1}\bar{B} = \begin{bmatrix} C(A+GC)^{k-1}(B+GD) & -C(A+GC)^{k-1}G \end{bmatrix}$ $\triangleq \begin{bmatrix} \bar{Y}_k^{(1)} & -\bar{Y}_k^{(2)} \end{bmatrix}$, and $Y_0 = D,\ Y_k = CA^{k-1}B$. $\because Y_1 = \bar{Y}_1^{(1)} - \bar{Y}_1^{(2)}D$, $Y_2 = \bar{Y}_2^{(1)} - \bar{Y}_2^{(2)}D - \bar{Y}_1^{(2)}Y_1,\ Y_3 = \bar{Y}_3^{(1)} - \bar{Y}_3^{(2)}D - \bar{Y}_1^{(2)}Y_2 - \bar{Y}_2^{(2)}Y_1, \cdots$. So, we can simplify them and get $D = Y_0 = \bar{Y}_0$, $\forall k = 1 \sim p,\ Y_k = \bar{Y}_k^{(1)} - \sum_{i=1}^{k} \bar{Y}_i^{(2)} Y_{k-i}$, $\forall k = p+1 \sim \infty,\ Y_k = -\sum_{i=1}^{p} \bar{Y}_i^{(2)} Y_{k-i}$. (Please refer to page 185 of textbook).

- $\Rightarrow \forall Y_1 \sim Y_{p+1}$, $\begin{bmatrix} I & 0 & \cdots & \cdots & 0 \\ \bar{Y}_1^{(2)} & I & \ddots & & \vdots \\ \bar{Y}_2^{(2)} & \bar{Y}_1^{(2)} & I & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \bar{Y}_p^{(2)} & \bar{Y}_{p-1}^{(2)} & \cdots & \bar{Y}_1^{(2)} & I \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \\ Y_{p+1} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1^{(1)} - \bar{Y}_1^{(2)}D \\ \bar{Y}_2^{(1)} - \bar{Y}_2^{(2)}D \\ \vdots \\ \bar{Y}_p^{(1)} - \bar{Y}_p^{(2)}D \\ 0 \end{bmatrix}$.

$\Rightarrow \forall k = p+2 \sim N \to \infty$, $\underbrace{\begin{bmatrix} -\bar{Y}_p^{(2)} & -\bar{Y}_{p-1}^{(2)} & \cdots & -\bar{Y}_1^{(2)} \end{bmatrix}}_{\triangleq \bar{\bar{y}}^{(2)}_{m\times mp}} \underbrace{\begin{bmatrix} Y_2 & Y_3 & \cdots & Y_{N+1} \\ Y_3 & Y_4 & \cdots & Y_{N+2} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{p+1} & Y_{p+2} & \cdots & Y_{p+N} \end{bmatrix}}_{=H(1)_{pm\times Nr}}$ $= \underbrace{\begin{bmatrix} Y_{p+2} & Y_{p+3} & \cdots & Y_{p+N+1} \end{bmatrix}}_{\triangleq \underline{Y}_{m\times rN}} \Rightarrow \bar{\bar{y}}^{(2)} H(1) = \underline{Y}$, this mapping is from observer *Markov* space; mp to the system *Markov* space; Nr after p^{th} steps.

- Because $\bar{Y}_k = \begin{bmatrix} \bar{Y}_k^{(1)} & -\bar{Y}_k^{(2)} \end{bmatrix} = 0,\ \forall k > p$ for FIR model, the maximum degree of freedom of $\begin{bmatrix} -\bar{Y}_p^{(2)} & -\bar{Y}_{p-1}^{(2)} & \cdots & -\bar{Y}_1^{(2)} \end{bmatrix}$ is mp. However, we need the $rank\{A\} = n$, it's the order of controllable and observable part of system. The maximum degree of $H(1)$ is $\min\{mp,\ Nr\}$ and maximum degree of freedom of $\begin{bmatrix} Y_{p+1} & Y_{p+2} & \cdots & Y_{p+N} \end{bmatrix}$ is $\min\{mp,\ n\} \geq n$. $\Rightarrow$ To keep the DOF, we need; $Nr \geq mp \geq n$ and to save computation, we keep p as small as possible. Please note that we must calculate until $Nr \geq mp$ and Y_{p+N+1}.

7.3 Observer gain *Markov* parameters:

- Find G from given observer *Markov* parameters: $\bar{Y}_k^{(1)}$ and $\bar{Y}_k^{(2)}$:

➢ Def: $Y_k^0 \triangleq CA^{k-1}G \ni H^0 = \begin{bmatrix} Y_1^0 & Y_2^0 & \cdots \\ Y_2^0 & Y_3^0 & \cdots \\ \vdots & \vdots & \\ Y_{k+1}^0 & Y_{k+2}^0 & \cdots \end{bmatrix} = \begin{bmatrix} CG & CAG & \cdots \\ CAG & CA^2G & \cdots \\ \vdots & \vdots & \\ CA^kG & CA^{k+1}G & \cdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^k \end{bmatrix}$

$\bullet \begin{bmatrix} G & AG & A^2G & \cdots \end{bmatrix}$, for solving G by ERA method.

➢ Relationship between Y_k^0 and $\bar{Y}_k^{(1)}$, $\bar{Y}_k^{(2)}$: comparing the definitions of them. We get $Y_1^0 = CG = \bar{Y}_1^{(2)}$, $\forall k = 2 \sim p$, $Y_k^0 = \bar{Y}_k^{(2)} - \sum_{i=1}^{k-1} \bar{Y}_i^{(2)} Y_{k-i}^0$,

and $\forall k = p+1 \sim \infty$, $Y_k^0 = -\sum_{i=1}^{p} \bar{Y}_i^{(2)} Y_{k-i}^0$ or (Please refer to page 188 of textbook)

$$\begin{bmatrix} I & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \bar{Y}_1^{(2)} & I & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \bar{Y}_{p-1}^{(2)} & \cdots & \bar{Y}_1^{(2)} & I & \ddots & & \vdots \\ \bar{Y}_p^{(2)} & \bar{Y}_{p-1}^{(2)} & \cdots & \bar{Y}_1^{(2)} & I & \ddots & \vdots \\ 0 & \bar{Y}_p^{(2)} & \ddots & \cdots & \bar{Y}_1^{(2)} & I & \\ \vdots & \ddots & \ddots & \ddots & \cdots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} Y_1^0 \\ Y_2^0 \\ \vdots \\ Y_p^0 \\ Y_{p+1}^0 \\ Y_{p+2}^0 \\ \vdots \end{bmatrix} = \begin{bmatrix} \bar{Y}_1^{(2)} \\ \bar{Y}_2^{(2)} \\ \vdots \\ \bar{Y}_p^{(2)} \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

Every Y_{k+1}^0 is a linear combination of previous $Y_k^0, \cdots, Y_{k-p}^0$ vectors.

➢ For solving G by ERA H^0: to match with ERA of H resulting $\begin{bmatrix} A & B & C \end{bmatrix}$, we use $\mathcal{Y}^0 \triangleq \begin{bmatrix} Y_1^0 \\ Y_2^0 \\ \vdots \\ Y_{k+1}^0 \end{bmatrix} = \begin{bmatrix} CG \\ CAG \\ \vdots \\ CA^kG \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^k \end{bmatrix} G \triangleq PG \Rightarrow$

$G = \left(P^T P\right)^{-1} P^T \mathcal{Y}^0$.

● Join estimation of system parameter and observer gain:

➢ Def: $\varGamma_k \triangleq \begin{bmatrix} Y_k & Y_k^0 \end{bmatrix} = \begin{bmatrix} CA^{k-1}B & CA^{k-1}G \end{bmatrix} = CA^{k-1}\begin{bmatrix} B & G \end{bmatrix} \triangleq CA^{k-1}B_C$.

Calculation of $\varGamma_k = \begin{bmatrix} \bar{Y}_k^{(1)} - \bar{Y}_k^{(2)}D & \bar{Y}_k^{(2)} \end{bmatrix} - \sum_{i=1}^{k-1} \bar{Y}_i^{(2)} \begin{bmatrix} Y_{k-i} & Y_{k-i}^0 \end{bmatrix}$

$= \begin{bmatrix} \bar{Y}_k^{(1)} - \bar{Y}_k^{(2)}D & \bar{Y}_k^{(2)} \end{bmatrix} - \sum_{i=1}^{k-1} \bar{Y}_i^{(2)} \varGamma_{k-i}$, $\forall k = 2 \sim p$ and $\varGamma_k = -\sum_{i=1}^{p} \bar{Y}_i^{(2)} \varGamma_{k-i}$, $\forall k = p+1 \sim \infty$. For $k=1$, $\varGamma_1 \triangleq \begin{bmatrix} Y_1 & Y_1^0 \end{bmatrix} = \begin{bmatrix} CB & CG \end{bmatrix}$.

➢ Def: $H_p = \begin{bmatrix} \varGamma_1 & \varGamma_2 & \cdots \\ \varGamma_2 & \varGamma_3 & \cdots \\ \vdots & \vdots & \\ \varGamma_k & \varGamma_{k+1} & \cdots \end{bmatrix} = \begin{bmatrix} CB_C & CAB_C & \cdots \\ CAB_C & CA^2B_C & \cdots \\ \vdots & \vdots & \\ CA^{k-1}B_C & CA^kB_C & \cdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}$

$\bullet \begin{bmatrix} B_C & AB_C & A^2B_C & \cdots \end{bmatrix}$. We can ERA H_p to find $\begin{bmatrix} A & B & C & G \end{bmatrix}$ di- rectly.

➢ Several different approach to find $\begin{bmatrix} A & B & C & G \end{bmatrix}$:

1. What's the difference between given model $y = ax$ to find a with $\min_a \left\| y - ax \right\|^2$ and $\min_{a^2} \sqrt{\left\| y^2 - a^2x^2 \right\|}$? The former result is optimal solution and the rear one is suboptimal solution.
2. ERA $\bar{Y}$ to find $\begin{bmatrix} \bar{A} & \bar{B} & C \end{bmatrix}$, then solve $\bar{B} = \begin{bmatrix} B + GD & -G \end{bmatrix}$ and $\bar{A} = A + GC$ for $\begin{bmatrix} A & B & C & G \end{bmatrix}$. The result is suboptimal for A, B.
3. ERA Y to find $\begin{bmatrix} A & B & C \end{bmatrix}$, then solve $\left(P^T P\right)^{-1} P^T \mathcal{Y}^0$ for G. The

result is suboptimal for G.

4. ERA H_p to find $\begin{bmatrix} A & B & C & G \end{bmatrix}$. This result is optimal for all of them.

If there is no noise then every approach gives same result. If noise exists then different approach has different object function therefore they have different results.

7.4 Relationship to *Kalman* filter:

- Noise to state space model; (Stochastic model): process noise $\omega(k)$, measurement noise $\upsilon(k)$. $\begin{matrix} x(k+1) = Ax(k) + Bu(k) + \omega(k) \\ y(k) = Cx(k) + Du(k) + \upsilon(k) \end{matrix}$.
- *Kalman* filter: it's designed with the knowledge of $\begin{bmatrix} A & B & C & G \end{bmatrix}$ and Q, R.
- Review *Kalman* filter: given state space system $\begin{bmatrix} A & B & C & G \end{bmatrix}$ and noise covariance to find a recursive state estimator; $\hat{x}(k+1) = F(k)\hat{x}(k) + K(k)y(k) + H(k)u(k)$ as update guess $\hat{x}(k)$ at k with modification according to output $y(k)$ and input $u(k)$.

➢ For the estimator to be good one, we define state estimation error $e(k) \square\, x(k) - \hat{x}(k)$. $e(k)$ will be a stochastic number because the existing noise $\omega(k)$ and $\upsilon(k)$.

➢ For a good estimation, $e(k)$ must be zero mean and have small covariance.

➢ Forming $e(k)$ as a function of $\begin{bmatrix} A & B & C & D \end{bmatrix}$, $\begin{bmatrix} \omega & \upsilon & \hat{x} & u \end{bmatrix}$, and $\begin{bmatrix} F & K & H \end{bmatrix}$: $e(k+1) \square\, x(k+1) - \hat{x}(k+1) = Ax(k) + Bu(k) + \omega(k)$
$-F(k)\hat{x}(k) - K(k)y(k) - H(k)u(k) = Ax(k) + Bu(k) + \omega(k)$

$-F(k)\hat{x}(k)-K(k)Cx(k)-K(k)Du(k)-K(k)\upsilon(k)-H(k)u(k)$

$=\left[A-K(k)C\right]\left[x(k)-\hat{x}(k)\right]+\left[A-K(k)C-F(k)\right]\hat{x}(k)-K(k)\upsilon(k)$

$+\left[B-K(k)D-H(k)\right]u(k)+\omega(k)\cdots(1).$

➢ Make $e(k)$ to be zero mean: $E\{e(k+1)\}=\left[A-K(k)C\right]E\{e(k)\}$ $+\left[A-K(k)C-F(k)\right]E\{\hat{x}(k)\}+\left[B-K(k)D-H(k)\right]E\{u(k)\}$ $-K(k)E\{\upsilon(k)\}+E\{\omega(k)\}=0$, where $\hat{x}(k)$ and $u(k)$ are determinist numbers $\Rightarrow E\{\hat{x}(k)\}=\hat{x}(k)$ and $E\{u(k)\}=u(k)$. The noises $\omega(k)$ and $\upsilon(k)$ are zero mean $\Rightarrow E\{\omega(k)\}=0$ and $E\{\upsilon(k)\}=0$. Thus far, $E\{e(k+1)\}=0,\ \Rightarrow A-K(k)C=F(k)\cdots(2)$, $B-K(k)D=H(k)\cdots(3)$, and $E\{e(k+1)\}=\left[A-K(k)C\right]E\{e(k)\}$

$=\left[A-K(k)C\right]\left[A-K(k-1)C\right]E\{e(k-1)\}$

$=\left[A-K(k)C\right]\left[A-K(k-1)C\right]\cdots E\{e(0)\}$. We need also $E\{e(0)\}=0$ which is $E\{x(0)-\hat{x}(0)\}=0\Rightarrow E\{\hat{x}(0)\}=E\{x(0)\}\triangleq\bar{x}(0)$. From Eqs. (2), (3) $\Rightarrow \hat{x}(k+1)=\left[A-K(k)C\right]\hat{x}(k)+\left[B-K(k)D\right]u(k)+K(k)y(k)$ compare with $\hat{x}(k+1)=(A+GC)\hat{x}(k)+(B+GD)u(k)-Gy(k)$; (page 3-2), we can find that $G=-K(k)$. It means that the G of OKID is the *Kalman* filter gain for zero mean estimation error.

● To calculate the *Kalman* filter gain $K(k)$ for R, Q, A, B, C to minimize the covariance $E\{e(k)e^T(k)\}$.

➢ From (1), (2), (3) $\Rightarrow e(k+1)=\left[A-K(k)C\right]e(k)-K(k)\upsilon(k)+\omega(k)$.

We define $\mathrm{P}(k+1)\triangleq E\{e(k+1)e^T(k+1)\}=\left[A-K(k)C\right]E\{e(k)e^T(k)\}$

$\bullet\left[A-K(k)C\right]^T+K(k)E\{\upsilon(k)\upsilon^T(k)\}K^T(k)+E\{\omega(k)\omega^T(k)\}$

$=\left[A-K(k)C\right]\mathrm{P}(k)\left[A-K(k)C\right]^T+K(k)RK^T(k)+Q\cdots(4)$, where

$E\left\{\upsilon(k)\upsilon^T(k)\right\} \triangleq R$, $E\left\{\omega(k)\omega^T(k)\right\} \triangleq Q$ and using $E\left\{e(k)\upsilon^T(k)\right\} = 0$, $E\left\{e(k)\omega^T(k)\right\} = 0$, $E\left\{\omega(k)\upsilon^T(k)\right\} \triangleq 0$ because current error only depends on previous input, process noise and measurement noise.

- To minimize $\mathbf{P}(k+1)$ with $K(k)$; $\dfrac{\partial \mathbf{P}(k+1)}{\partial K(k)} = 0$:

$\Rightarrow -2\left[A - K(k)C\right]\mathbf{P}(k)C^T + 2K(k)R = 0$

$\Rightarrow K(k) = A\mathbf{P}(k)C^T\left[R + C\mathbf{P}(k)C^T\right]^{-1}$; it's the so-call optimal gain or *Kalman* gain. Then, we substitute $K(k)$ into the formula of $\mathbf{P}(k+1)$;

$\Rightarrow \mathbf{P}(k+1) = A\mathbf{P}(k)A^T - A\mathbf{P}(k)C^T\left[R + C\mathbf{P}(k)C^T\right]^{-1}C\mathbf{P}(k)A^T + Q$; it's known as discrete algebraic *Riccati* equation.

- Steady-state *Kalman* filter:

- For steady state, $K(k+1) = K(k) = \cdots = K$ and $\mathbf{P}(k+1) = \mathbf{P}(k) = \cdots = \mathbf{P}$, $\Rightarrow K = A\mathbf{P}C^T\left[R + C\mathbf{P}C^T\right]^{-1}$ and $\mathbf{P} = A\mathbf{P}A^T - A\mathbf{P}C^T\left[R + C\mathbf{P}C^T\right]^{-1}C\mathbf{P}A^T + Q$. Process noise $\omega(k)$ covariance keeps error covariance large and measurement noise $\upsilon(k)$ keep error covariance large, too. Because $R\uparrow \Rightarrow \left[R + C\mathbf{P}C^T\right]^{-1}\downarrow \Rightarrow \mathbf{P}\uparrow$.

- Statistic characteristics of $e(k)$:

1. The mean of $e(k)$: for $G = -K$; *Kalman* gain, $E\left\{e(k)\right\} = 0$.
2. The covariance of $e(k)$: from the above, we know $E\left\{e(k+1)e^T(k)\right\}$
   $= E\left\{\left[(A-KC)e(k) - K\upsilon(k) + \omega(k)\right]e^T(k)\right\} = (A-KC)E\left\{e(k)e^T(k)\right\}$
   $-KE\left\{\upsilon(k)e^T(k)\right\} + E\left\{\omega(k)e^T(k)\right\} = (A-KC)\mathbf{P}$

   $\Rightarrow E\left\{e(k+\tau)e^T(k)\right\} = (A-KC)^\tau\,\mathbf{P}$.

   [Pf]: for example, $E\left\{e(k+2)e^T(k)\right\}$; $\because e(k+2)e^T(k)$

$$= e(k+2)e^T(k+1)\left[e^T(k+1)\right]^{\#}\left[e(k+1)\right]^{\#} e(k+1)e^T(k)$$

$$= e(k+2)e^T(k+1)\left[e(k+1)e^T(k+1)\right]^{\#} e(k+1)e^T(k)$$

$$\therefore E\begin{Bmatrix} e(k+2) \\ e^T(k) \end{Bmatrix} = E\begin{Bmatrix} e(k+2) \\ e^T(k+1) \end{Bmatrix}\left[E\begin{Bmatrix} e(k+1) \\ e^T(k+1) \end{Bmatrix}\right]^{\#} E\begin{Bmatrix} e(k+1) \\ e^T(k) \end{Bmatrix}$$

$$= (A-KC)\mathrm{P}\mathrm{P}^{\#}(A-KC)\mathrm{P} = (A-KC)^2\mathrm{P}.$$

3. $e(k)$ is zero mean but not white.

7.5 Residual of output:

- Def: $y(k) = C\hat{x}(k) + Du(k) + \varepsilon_r(k)$, $\hat{x}(k+1) = \tilde{A}\hat{x}(k) + \tilde{B}\bar{u}(k)$,

where $\tilde{A} = A - KC$, $\tilde{B} = \begin{bmatrix} B - KD & K \end{bmatrix}$ and $\bar{u}(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}$.

- Mean of $\varepsilon_r(k)$:

$$\because y(k) = Cx(k) + Du(k) + \upsilon(k) = C\hat{x}(k) + Du(k) + \varepsilon_r(k)$$

$$\Rightarrow C\left[x(k) - \hat{x}(k)\right] + \upsilon(k) = Ce(k) + \upsilon(k) = \varepsilon_r(k)$$

$$\Rightarrow E\left\{\varepsilon_r(k)\right\} = CE\left\{e(k)\right\} + E\left\{\upsilon(k)\right\} = 0.$$

- Covariance of $\varepsilon_r(k)$:

$$E\left\{\varepsilon_r(k+1)\varepsilon_r^T(k)\right\} = E\left\{\left[Ce(k+1) + \upsilon(k+1)\right]\left[Ce(k) + \upsilon(k)\right]^T\right\}$$

$$= CE\left\{e(k+1)e^T(k)\right\}C^T + CE\left\{e(k+1)\upsilon^T(k)\right\} + E\left\{\upsilon(k+1)e^T(k)\right\}C^T$$

$$+E\left\{\upsilon(k+1)\upsilon^T(k)\right\} = C(A-KC)\mathrm{P}C^T + C(A-KC)\cancel{E\left\{e(k)\upsilon^T(k)\right\}}^{0}$$

$$-CKE\left\{\upsilon(k)\upsilon^T(k)\right\} + C\cancel{E\left\{\omega(k)\upsilon^T(k)\right\}}^{0} + \cancel{E\left\{\upsilon(k+1)e^T(k)\right\}}^{0}C^T$$

$$+E\cancel{\left\{\upsilon(k+1)\upsilon^T(k)\right\}}^{0};\ \because e(k+1) = \left[A - K(k)C\right]e(k) - K(k)\upsilon(k) + \omega(k)$$

$$= CA\mathrm{P}C^T - CKC\mathrm{P}C^T - CKR = CA\mathrm{P}C^T - CK\left(R + C\mathrm{P}C^T\right) = CA\mathrm{P}C^T$$

$$-C\left[A\mathrm{P}C^T\left(R + C\mathrm{P}C^T\right)^{-1}\right]\left(R + C\mathrm{P}C^T\right);\ \because K = A\mathrm{P}C^T\left[R + C\mathrm{P}C^T\right]^{-1}$$

$= CAPC^T - CAPC^T = 0$. Residual information in output signal is zero mean, white *Gaussian* for *Kalman* filter therefore *Kalman* filter is also named Whiten filter.

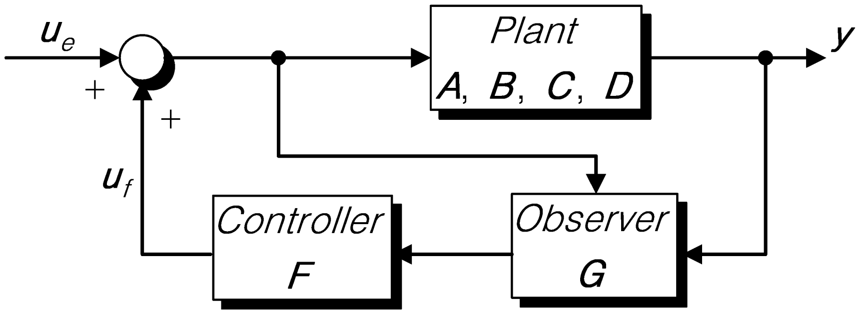
- Homework 11:

1. Simulate the following on the previous spring-mass-damper system: please add a little white noise into the one of all inputs and the one of all outputs, respectively.
2. Take RLS on the result of the above simulation to obtain the *Markov* parameters.
3. Use the OKID method to identify the system parameters and optimal *Kalman* gain.
4. Compare the results from different noise levels.
5. Check the residual of output whether white noise or not.
6. Compare the G from OKID with the *Kalman* gain (obtained from *Riccatic* equation.

8 Observer / Controller identification; OCID:

8.1 System description:

- A closed loop system under identification:



- What's OCID?

$$\left.\begin{array}{l} u_e \\ u_f \\ y \end{array}\right. \xrightarrow{LS} \begin{array}{l} Closed\ Loop \\ Observer\ Markov \\ parameters \end{array} \rightarrow \left.\begin{array}{l} Open\ system \\ Observer \\ Controller \end{array}\right\} Markov\ parameters$$

$$\xrightarrow{ERA} \begin{array}{l} Open\ system\ model \\ Observer\ gain \\ Controller\ gain \end{array}$$

. It's an extension of OKID method to ID a close loop system.

- When does one use OCID?

When systems are inherently unstable, only closed loop data are available for identification. This method can obtain system model and control gain (true gain).

- Models:

➢ Dynamic system: $\begin{array}{l} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{array}$.

➢ Observer: $\begin{array}{l} \bar{x}(k+1) = A\bar{x}(k) + Bu(k) - G_e\left[y(k) - \bar{y}(k)\right] \\ \bar{y}(k) = C\bar{x}(k) + Du(k) \end{array}$.

➢ State feedback control with excitation: $u(k) = u_f(k) + u_e(k)$, where $u_f(k) = -F\bar{x}(k)$ and $u_e(k)$ is extra for butter ID excitation.

➢ G_e and $-F$: they are designed by engineer but doesn't represent the true value and need to be identified due to the variation of components.

➢ $u_e(k)$: it is extra for butter ID excitation and used to avoid linearly depends of $y(k)$ and $u(k)$ because it will defect the rank of $\bar{u}(k)$.

- Summary of models:

$$\bar{x}(k+1) = (A+G_eC)\bar{x}(k) + (B+G_eD)u(k) - G_e y(k)$$

$$= (A+G_eC)\bar{x}(k) + \begin{bmatrix} B+G_eD & -G_e \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \triangleq \bar{A}_e\bar{x}(k) + \bar{B}_e\bar{u}(k),$$

$$\begin{bmatrix} \bar{y}(k) \\ u_f(k) \end{bmatrix} = \begin{bmatrix} C \\ -F \end{bmatrix}\bar{x}(k) + \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \triangleq \bar{C}\bar{x}(k) + \bar{D}\bar{u}(k).$$ Thus far, the state, input and output of the new model are $\bar{x}(k)$, $\bar{u}(k)$ and $\begin{bmatrix} \bar{y}(k) \\ u_f(k) \end{bmatrix}$, respectively.

8.2 *Markov* model of closed loop Observer:

- Just for stable observer and finite $\hat{k}$; (FIR model), then

$$\begin{bmatrix} \bar{y}(k) \\ u_f(k) \end{bmatrix} = \sum_{i=1}^{\hat{k}} \bar{Y}_i\bar{u}(k-i) + \bar{Y}_0\bar{u}(k), \text{ where } \begin{matrix} \bar{Y}_i = \bar{C}\bar{A}_e^{i-1}\bar{B}_e, \ \forall i = 1 \sim \hat{k} \\ \bar{Y}_0 = \bar{D} \end{matrix}.$$

➢ We use ERA to construct $\bar{A}_e$, $\bar{B}_e$, $\bar{C}$, $\bar{D}$.

➢ How to estimate $\bar{Y}_e$; for $\ell > \hat{k}$: $\bar{y}_e \triangleq \begin{bmatrix} \bar{y}(0) & \bar{y}(1) & \cdots & \bar{y}(\ell) \\ u_f(0) & u_f(1) & \cdots & u_f(\ell) \end{bmatrix}$

$$= \begin{bmatrix} \bar{D} & \bar{C}\bar{B}_e & \cdots & \bar{C}\bar{A}_e^{\hat{k}-1}\bar{B}_e \end{bmatrix}\begin{bmatrix} \bar{u}(0) & \bar{u}(1) & \cdots & \bar{u}(\hat{k}) & \cdots & \bar{u}(\ell) \\ 0 & \bar{u}(0) & \cdots & \bar{u}(\hat{k}-1) & \cdots & \bar{u}(\ell-1) \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & \bar{u}(0) & \cdots & \bar{u}(\ell-\hat{k}) \end{bmatrix}$$

$\triangleq \bar{Y}_e\mathcal{U}_p$. We can use LS to obtain $\Rightarrow \bar{Y}_e = \bar{y}_e\mathcal{U}_p^T\left(\mathcal{U}_p\mathcal{U}_p^T\right)^{-1}$.

- Problem of above process: Although the $u_f(k)$ and $\bar{u}(k) \triangleq \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}$ generated by controller or measurement data are all available to us, $\bar{y}(k) = C\bar{x}(k) + Du(k)$ is estimated base on $\begin{bmatrix} A & B & C & G_e \end{bmatrix}$ and $\bar{x}(k)$ which will be identified. Therefore, $\bar{y}(k)$ isn't available.

- If observer is stable. $\Rightarrow$ for large $k \Rightarrow \bar{x}(k) \to x(k) \Rightarrow \bar{y}(k) \to y(k)$.

This means that $\bar{y}(k)$ can be replaced by $y(k)$ for large k and we may take data at large k only for LS estimation. In this time, the system can be rewritten as $\begin{array}{l} u_f(k) = -Fx(k) \\ u(k) = u_f(k) + u_e(k) \end{array}$, $\begin{array}{l} \bar{y}(k) = y(k) \\ = Cx(k) + Du(k) \end{array}$ and

$$x(k+1) = Ax(k) + Bu(k) + Gy(k) - Gy(k) = (A+GC)x(k)$$

$$+(B+GD)u(k) - Gy(k) = (A+GC)x(k) + \begin{bmatrix} B+GD & -G \end{bmatrix} \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}$$

$$\triangleq \bar{A}x(k) + \bar{B}\bar{u}(k) \Rightarrow \begin{bmatrix} y(k) \\ u_f(k) \end{bmatrix} = \begin{bmatrix} C \\ -F \end{bmatrix} x(k) + \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}$$

$$\triangleq \bar{C}x(k) + \bar{D}\bar{u}(k) \Rightarrow y_t \triangleq \begin{bmatrix} y(k) & y(k+1) & \cdots & y(\ell) \\ u_f(k) & u_f(k+1) & \cdots & u_f(\ell) \end{bmatrix}$$

$$= \begin{bmatrix} \bar{D} & \bar{C}\bar{B} & \cdots & \bar{C}\bar{A}^{\hat{k}-1}\bar{B} \end{bmatrix} \begin{bmatrix} \bar{u}(k) & \cdots & \bar{u}(\hat{k}) & \cdots & \bar{u}(\ell) \\ \vdots & & \vdots & & \vdots \\ \bar{u}(0) & \cdots & \bar{u}(\hat{k}-k) & \cdots & \bar{u}(\ell-k) \\ 0 & \ddots & \vdots & & \vdots \\ 0 & 0 & \bar{u}(0) & \cdots & \bar{u}(\ell-\hat{k}) \end{bmatrix} \triangleq \bar{Y}\mathcal{U}_t.$$

Therefore, we can use LS to obtain $\Rightarrow \bar{Y} = y_t \mathcal{U}_t^T \left(\mathcal{U}_t \mathcal{U}_t^T\right)^{-1} = y_t \mathcal{U}_t^{\#}$. The estimation of G is different from G_e and is optimal gain instead.

8.3 System and Gain *Markov* parameters:

- Given $\bar{Y} \xrightarrow{ERA} \begin{array}{ll} \bar{A} = A + GC & \bar{B} = \begin{bmatrix} B+GD & -G \end{bmatrix} \\ \bar{C} = \begin{bmatrix} C \\ -F \end{bmatrix} & \bar{D} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \end{array} \xrightarrow{solve} \begin{array}{ll} A & B \\ C & D \\ F & G \end{array}$,

where $C \quad D \quad F \quad G$ are optimal and $A \quad B$ are suboptimal.

- For optimal in A, B, C, D, F, G, we need to combine state space model such that A, B, C, D, F, G can be solved simultaneously from

Markov parameters. First, we define some *Markov* parameters as follow;

$$Y_k \triangleq \begin{bmatrix} C \\ F \end{bmatrix} A^{k-1} \begin{bmatrix} B & G \end{bmatrix} = \begin{bmatrix} CA^{k-1}B & CA^{k-1}G \\ FA^{k-1}B & FA^{k-1}G \end{bmatrix} \triangleq \begin{bmatrix} Y_k(1,1) & Y_k(1,2) \\ Y_k(2,1) & Y_k(2,2) \end{bmatrix},$$

$\forall k = 1,2,\cdots$ and $Y_0(1,1) = D$. From the page 3-4 and 3-5, we find that the element $Y_k(1,1) = CA^{k-1}B$ is the system *Markov* parameters and $Y_k(1,2) = CA^{k-1}G$ is the observer gain *Markov* parameters. In addition, $Y_k(2,1) = FA^{k-1}B$ and $Y_k(2,2) = FA^{k-1}G$ are controller gain and observer / controller gain *Markov* parameters, respectively. However, the *Markov* parameters of the steady state summary model on page 3-16 is

$\forall k = 1,2,\cdots$, $\bar{Y}_k = \bar{C}\bar{A}^{k-1}\bar{B} = \begin{bmatrix} C \\ -F \end{bmatrix} (A+GC)^{k-1} \begin{bmatrix} B+GD & -G \end{bmatrix}$

$= \begin{bmatrix} C(A+GC)^{k-1}(B+GD) & -C(A+GC)^{k-1}G \\ -F(A+GC)^{k-1}(B+GD) & F(A+GC)^{k-1}G \end{bmatrix} \triangleq \begin{bmatrix} \bar{Y}_k(1,1) & -\bar{Y}_k(1,2) \\ -\bar{Y}_k(2,1) & \bar{Y}_k(2,2) \end{bmatrix}$ and

$\bar{Y}_0 = \begin{bmatrix} \bar{Y}_0(1,1) & -\bar{Y}_0(1,2) \\ -\bar{Y}_0(2,1) & \bar{Y}_0(2,2) \end{bmatrix} = \bar{D} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$.

- Second, by the same manner described on page 3-4, $\forall k = 1 \sim \infty$,

$Y_k(1,1) = \bar{Y}_k(1,1) - \sum_{i=1}^{k} \bar{Y}_i(1,2) Y_{k-i}(1,1)$, $Y_k(1,2) = \bar{Y}_k(1,2) - \sum_{i=1}^{k-1} \bar{Y}_i(1,2) Y_{k-i}(1,2)$

but $\forall k > p$, $\bar{Y}_k(1,1) = \bar{Y}_k(1,2) = 0$. As regards the $Y_k(2,1) = FA^{k-1}B$, we take $k=3$ for an example. $\bar{Y}_3(2,1) = F(A+GC)^2(B+GD) = FA^2B + FGCAB$ $+F(A+GC)GCB + F(A+GC)^2 GD = Y_3(2,1) + \bar{Y}_1(2,2)Y_2(1,1) + \bar{Y}_2(2,2)Y_1(1,1)$ $+\bar{Y}_3(2,2)Y_0(1,1) \Rightarrow Y_3(2,1) = \bar{Y}_3(2,1) - \bar{Y}_1(2,2)Y_2(1,1) - \bar{Y}_2(2,2)Y_1(1,1) - \bar{Y}_3(2,2)Y_0(1,1)$. In general, $\forall k = 1 \sim \infty$, $Y_k(2,1) = \bar{Y}_k(2,1) - \sum_{i=1}^{k} \bar{Y}_i(2,2) Y_{k-i}(1,1)$. In the same ideal, $Y_k(2,2) = \bar{Y}_k(2,2) - \sum_{i=1}^{k} \bar{Y}_i(2,2) Y_{k-i}(1,2)$, where $\bar{Y}_k(2,1) = \bar{Y}_k(2,2) = 0$, $\forall k > p$.

- In summary, we can use these result to complete the process as follow; $\bar{Y}_k \xrightarrow{above} Y_k \xrightarrow{ERA} \begin{bmatrix} A & B & C & D & F & G \end{bmatrix}$; optimal solution.

9 Frequency-Domain State Space system ID:

9.1 Introduction:

- State space models are required for modern control design.
- OKID constructs state space models with time domain I/O data.
- There are cases in which frequency response data are available only.
- IDTFT transform frequency domain to time domain but with time aliasing effect.
- SSFD; State Space Frequency Domain:
  - Estimate *Markov* parameters from FRF.
  - Without distorting; (time aliasing effect).
  - Arbitrary frequency weighting can be introduced.
  - Curve-fitting rational matrix description; ratio of matrix polynomial and monic scalar polynomial denominator.
  - Nonlinear Solver or linear approximation with iteration.
- Method introduced in this book:
  - Curve-fitting Matrix-Fraction.
  - Linear least square solver.
- State space; $\begin{matrix} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{matrix}$, $x_{n\times 1}(k)$: state vector,

$u_{r\times 1}(k)$: input vector, $y_{m\times 1}(k)$: output vector, $A_{n\times n}(k)$: system matrix,

$B_{n\times r}(k)$: input matrix, $C_{m\times n}(k)$: output matrix, $D_{m\times r}(k)$: direct transmission matrix.

- FRF vs. S.S. : $G(z_k) = C(z_k I_n - A)^{-1} B + D$; $z_k = e^{j\frac{2\pi}{\ell}k}$, where ℓ is the length of data, z_k correspond to frequency $\dfrac{2\pi k}{\ell \Delta t}$.
- SSFD: for $G(z_k)$, find $\hat{A}, \hat{B}, \hat{C}, \hat{D} \ni \hat{G}(z_k) = \hat{C}(z_k I_n - \hat{A})^{-1} \hat{B} + \hat{D}$ and minimize $\mathcal{J} \triangleq \sum_{k=0}^{\ell-1} \omega^2(z_k) \left\| G(z_k) - \hat{G}(z_k) \right\|_2$, where $\omega(z_k)$ is a specified frequency weighting function.

➢ Nonlinear optimal problem.

➢ Separate to two steps:

1. Optimal *Markov* parameters.
2. ERA.

➢ Take $\tilde{G}(z_k) = \hat{D} + \hat{C}\hat{B}z_k^{-1} + \hat{C}\hat{A}\hat{B}z_k^{-2} + \hat{C}\hat{A}^2\hat{B}z_k^{-3} + \cdots = \sum_{i=0}^{\infty} \hat{Y}_i z_k^{-i}$.

9.2 Linear Curve-Fitting *Markov*:

- $\min_{\hat{Y}_k} \mathcal{J} \triangleq \min_{\hat{Y}_k} \sum_{k=0}^{\ell-1} \omega^2(z_k) \left\| G(z_k) - \sum_{i=0}^{\infty} \hat{Y}_i z_k^{-i} \right\|_2$. In theorem, there is an infinite term. In real case, it's very long even after truncation.
- $\hat{G}(z_k) = \hat{C}(z_k I_n - \hat{A})^{-1} \hat{B} + \hat{D} = \dfrac{1}{d(z_k)} \left[\hat{C}\mathrm{adj}(z_k I_n - \hat{A}) \hat{B} + d(z_k) \hat{D} \right]$.

It's non-unique (We can over specify the order of polynomial) and non-linear.

9.3 Left Matrix-Fraction Description; LMFD:

- Matrix Fraction:

Left: $\underset{m\times r}{G(z_k)} = \underset{m\times m}{\bar{Q}^{-1}(z_k)} \underset{m\times r}{\bar{R}(z_k)} \rightarrow$ observable canonical form.

Right: $\underset{m\times r}{G(z_k)} = \underset{m\times r}{R(z_k)}\underset{r\times r}{Q^{-1}(z_k)} \to$ controllable canonical form.

- LMFD: $G(z_k) = \bar{Q}^{-1}(z_k)\bar{R}(z_k)$, where

$\bar{Q}(z_k) = I_m + \bar{Q}_1 z_k^{-1} + \cdots + \bar{Q}_p z_k^{-p}$ and $\bar{R}(z_k) = \bar{R}_0 + \bar{R}_1 z_k^{-1} + \cdots + \bar{R}_p z_k^{-p}$.

They are non-unique. Review state space model under steady state;

$$x(k+1) = \bar{A}x(k) + \bar{B}\bar{u}(k) = (A+GC)x(k) + \begin{bmatrix} B+GD & -G \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \text{ and}$$

$$y(k) = Cx(k) + Du(k). \Rightarrow x(z) = \left(zI - \bar{A}\right)^{-1}\bar{B}\bar{u}(z) \text{ and}$$

$$y(z) = C\left(zI - \bar{A}\right)^{-1}\bar{B}\bar{u}(z) + Du(z) = C\left(zI - \bar{A}\right)^{-1}\begin{bmatrix} B+GD & -G \end{bmatrix}\begin{bmatrix} u(z) \\ y(z) \end{bmatrix}$$

$$+Du(z) = -C\left(zI - \bar{A}\right)^{-1}Gy(z) + \left[C\left(zI - \bar{A}\right)^{-1}(B+GD) + D\right]u(z)$$

$$\Rightarrow \left[I + C\left(zI - \bar{A}\right)^{-1}G\right]y(z) = \left[C\left(zI - \bar{A}\right)^{-1}(B+GD) + D\right]u(z)$$

$$\Rightarrow y(z) = G(z)u(z) = \bar{Q}^{-1}(z)\bar{R}(z)u(z) = \left[I + C\left(zI - \bar{A}\right)^{-1}G\right]^{-1}$$

$$\bullet\left[C\left(zI - \bar{A}\right)^{-1}(B+GD) + D\right]u(z),\ \therefore \bar{Q}(z) = I_m + \bar{Q}_1 z^{-1} + \cdots + \bar{Q}_p z^{-p}$$

$$= \left[I + C\left(zI - \bar{A}\right)^{-1}G\right]^{-1} = I + \bar{Y}_1^{(2)}z^{-1} + \bar{Y}_2^{(2)}z^{-2} + \cdots + \bar{Y}_p^{(2)}z^{-p} \text{ and}$$

$$\bar{R}(z) = \bar{R}_0 + \bar{R}_1 z^{-1} + \cdots + \bar{R}_p z^{-p} = C\left(zI - \bar{A}\right)^{-1}(B+GD) + D$$

$$= \bar{Y}_0^{(1)} + \bar{Y}_1^{(1)}z^{-1} + \bar{Y}_2^{(1)}z^{-2} + \cdots + \bar{Y}_p^{(1)}z^{-p}. \Rightarrow \bar{R}_0 = \bar{Y}_0^{(1)} = D, \cdots,$$

$$\bar{R}_p = \bar{Y}_p^{(1)} = C\bar{A}^{p-1}(B+GD) \text{ and } \bar{Q}_1 = \bar{Y}_1^{(2)} = CG, \cdots, \bar{Q}_p = \bar{Y}_p^{(2)} = C\bar{A}^{P-1}G.$$

$$\because G(z_k) = \bar{Q}^{-1}(z_k)\bar{R}(z_k) \Rightarrow \bar{Q}(z_k)G(z_k) = \bar{R}(z_k)$$

$$\Rightarrow \left(I_m + \bar{Q}_1 z_k^{-1} + \cdots + \bar{Q}_p z_k^{-p}\right)G(z_k) = \bar{R}_0 + \bar{R}_1 z_k^{-1} + \cdots + \bar{R}_p z_k^{-p}$$

$$\Rightarrow G(z_k) = -\bar{Q}_1 G(z_k)z_k^{-1} - \cdots - \bar{Q}_p G(z_k)z_k^{-p} + \bar{R}_0 + \bar{R}_1 z_k^{-1} + \cdots + \bar{R}_p z_k^{-p}$$

$$
=\begin{bmatrix} -\bar{Q}_1 & \cdots & -\bar{Q}_p & \bar{R}_0 & \bar{R}_1 & \cdots & \bar{R}_p \end{bmatrix}\begin{bmatrix} G(z_k)z_k^{-1} \\ \vdots \\ G(z_k)z_k^{-p} \\ 1 \\ z_k^{-1} \\ \vdots \\ z_k^{-p} \end{bmatrix},\ G(z_k) \text{ is known for}
$$

$z_k = e^{j\frac{2\pi}{\ell}k}, \forall k = 0, \cdots, \ell-1$. We collect them as matrix; $\bar{\Psi} = \bar{\Theta}\bar{\Phi}$ where

$$
\bar{\Phi} = \begin{bmatrix} G(z_0)z_0^{-1} & G(z_1)z_1^{-1} & \cdots & G(z_{\ell-1})z_{\ell-1}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ G(z_0)z_0^{-p} & G(z_1)z_1^{-p} & \cdots & G(z_{\ell-1})z_{\ell-1}^{-p} \\ I_r & I_r & \cdots & I_r \\ z_0^{-1}I_r & z_1^{-1}I_r & \cdots & z_{\ell-1}^{-1}I_r \\ \vdots & \vdots & \ddots & \vdots \\ z_0^{-p}I_r & z_1^{-p}I_r & \cdots & z_{\ell-1}^{-p}I_r \end{bmatrix},
$$

$\bar{\Theta} = \begin{bmatrix} -\bar{Q}_1 & \cdots & -\bar{Q}_p & \bar{R}_0 & \bar{R}_1 & \cdots & \bar{R}_p \end{bmatrix}$ and

$\bar{\Psi} = \begin{bmatrix} G(z_0) & G(z_1) & \cdots & G(z_{\ell-1}) \end{bmatrix}$.

- These are linear equations (complex number equations) where $\bar{\Theta}$ may be complex. If we force $\bar{\Theta}$ to be real then $\mathrm{Re}\{\bar{\Psi}\} = \bar{\Theta}\mathrm{Re}\{\bar{\Phi}\}$ and $\mathrm{Im}\{\bar{\Psi}\} = \bar{\Theta}\mathrm{Im}\{\bar{\Phi}\}$. Join together, we can obtain $\begin{bmatrix} \mathrm{Re}\{\bar{\Psi}\} \\ \mathrm{Im}\{\bar{\Psi}\} \end{bmatrix} = \bar{\Theta}\begin{bmatrix} \mathrm{Re}\{\bar{\Phi}\} \\ \mathrm{Im}\{\bar{\Phi}\} \end{bmatrix}$

$$
\Rightarrow \bar{\Theta} = \begin{bmatrix} \mathrm{Re}\{\bar{\Psi}\} \\ \mathrm{Im}\{\bar{\Psi}\} \end{bmatrix}\begin{bmatrix} \mathrm{Re}\{\bar{\Phi}\} \\ \mathrm{Im}\{\bar{\Phi}\} \end{bmatrix}^{\#}.
$$

- From page 3-4, it implies that the Estimation of system *Markov* parameters is $\Rightarrow Y_0 = D = \bar{R}_0,\ Y_k = \bar{R}_k - \sum_{i=1}^{k}\bar{Q}_i Y_{k-i}$, $\forall k = 1 \sim p$, and

$Y_k = -\sum_{i=1}^{k}\bar{Q}_i Y_{k-i}$, $\forall k > p$.

- Observable Canonical-Form realization: $\because y(z_k)=G(z_k)u(z_k)$

$=\bar{Q}^{-1}(z_k)\bar{R}(z_k)u(z_k)\Rightarrow \bar{Q}(z_k)y(z_k)=\bar{R}(z_k)u(z_k)$ and $\bar{R}_0=D$

$\Rightarrow \left(I_m+\bar{Q}_1 z_k^{-1}+\cdots+\bar{Q}_p z_k^{-p}\right)y(z_k)=\left(D+\bar{R}_1 z_k^{-1}+\cdots+\bar{R}_p z_k^{-p}\right)u(z_k)$,

$\Rightarrow \left[I_m y(z_k)-Du(z_k)\right]+\left[\bar{Q}_1 y(z_k)-\bar{R}_1 u(z_k)\right]z_k^{-1}+\cdots$

$+\left[\bar{Q}_p y(z_k)-\bar{R}_p u(z_k)\right]z_k^{-p}=0$. Then, we let $x_p(z_k)\triangleq y(z_k)-Du(z_k)$,

$x_{p-1}(z_k)\triangleq\left[I_m y(z_k)-Du(z_k)\right]z_k+\left[\bar{Q}_1 y(z_k)-\bar{R}_1 u(z_k)\right]=x_p(z_k)z_k$

$+\left[\bar{Q}_1 y(z_k)-\bar{R}_1 u(z_k)\right]$, $\cdots$, $x_1(z_k)\triangleq x_2(z_k)z_k+\left[\bar{Q}_{p-1}y(z_k)-\bar{R}_{p-1}u(z_k)\right]$

$\Rightarrow x_0(z_k)\triangleq x_1(z_k)z_k+\left[\bar{Q}_p y(z_k)-\bar{R}_p u(z_k)\right]=0$.

$\because y(z_k)=x_p(z_k)+Du(z_k)\Rightarrow x_{p-1}(z_k)=x_p(z_k)z_k+\bar{Q}_1 x_p(z_k)$

$+\left[\bar{Q}_1 D-\bar{R}_1\right]u(z_k)$, $\cdots$, $x_1(z_k)=x_2(z_k)z_k+\bar{Q}_{p-1}x_p(z_k)$

$+\left[\bar{Q}_{p-1}D-\bar{R}_{p-1}\right]u(z_k)$, $x_1(z_k)z_k=-\bar{Q}_p x_p(z_k)+\left[\bar{R}_p-\bar{Q}_p D\right]u(z_k)$,

$$\Rightarrow \begin{bmatrix} x_1\\ x_2\\ x_3\\ \vdots\\ x_p\end{bmatrix} z_k=\begin{bmatrix} 0 & \cdots & \cdots & 0 & -\bar{Q}_p\\ I & \ddots & & \vdots & -\bar{Q}_{p-1}\\ 0 & I & \ddots & \vdots & -\bar{Q}_{p-2}\\ \vdots & \ddots & \ddots & 0 & \vdots\\ 0 & \cdots & 0 & I & -\bar{Q}_1\end{bmatrix}\begin{bmatrix} x_1\\ x_2\\ x_3\\ \vdots\\ x_p\end{bmatrix}+\begin{bmatrix} \bar{R}_p-\bar{Q}_p D\\ \bar{R}_{p-1}-\bar{Q}_{p-1}D\\ \bar{R}_{p-2}-\bar{Q}_{p-2}D\\ \vdots\\ \bar{R}_1-\bar{Q}_1 D\end{bmatrix}u\,.$$

9.4 Right Matrix-Fraction Description; RMFD:

- RMFD: $G(z_k)=R(z_k)Q^{-1}(z_k)\Rightarrow G(z_k)Q(z_k)=R(z_k)$,

$\Rightarrow G(z_k)\left(I_r+Q_1 z_k^{-1}+\cdots+Q_p z_k^{-p}\right)=R_0+R_1 z_k^{-1}+\cdots+R_p z_k^{-p}$

$\Rightarrow G(z_k)=-G(z_k)z_k^{-1}Q_1-\cdots-G(z_k)z_k^{-p}Q_p+R_0+R_1 z_k^{-1}+\cdots+R_p z_k^{-p}$,

$G(z_{k-1})=-G(z_{k-1})z_{k-1}^{-1}Q_1-\cdots-G(z_{k-1})z_{k-1}^{-p}Q_p+R_0+R_1 z_{k-1}^{-1}+\cdots+R_p z_{k-1}^{-p}$,

$\cdots$, $G(z_0)=-G(z_0)z_0^{-1}Q_1-\cdots-G(z_0)z_0^{-p}Q_p+R_0+R_1 z_0^{-1}+\cdots+R_p z_0^{-p}$. We

collect them as matrix; $\Psi=\Phi\Theta$ where

$$\boldsymbol{\Phi}=\begin{bmatrix} G(z_0)z_0^{-1} & \cdots & G(z_0)z_0^{-p} & I_m & z_0^{-1}I_m & \cdots & z_0^{-p}I_m \\ G(z_1)z_1^{-1} & \cdots & G(z_1)z_1^{-p} & I_m & z_1^{-1}I_m & \cdots & z_1^{-p}I_m \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ G(z_{\ell-1})z_{\ell-1}^{-1} & \cdots & G(z_{\ell-1})z_{\ell-1}^{-p} & I_m & z_{\ell-1}^{-1}I_m & \cdots & z_{\ell-1}^{-p}I_m \end{bmatrix},\ \boldsymbol{\Theta}=\begin{bmatrix} -Q_1 \\ \vdots \\ -Q_p \\ R_0 \\ R_1 \\ \vdots \\ R_p \end{bmatrix}$$

and $\boldsymbol{\Psi}=\begin{bmatrix} G(z_0) & G(z_1) & \cdots & G(z_{\ell-1}) \end{bmatrix}^T$.

- If we force $\boldsymbol{\Theta}$ to be real then

$$\begin{bmatrix} \mathrm{Re}\{\boldsymbol{\Psi}\} \\ \mathrm{Im}\{\boldsymbol{\Psi}\} \end{bmatrix}=\begin{bmatrix} \mathrm{Re}\{\boldsymbol{\Phi}\} \\ \mathrm{Im}\{\boldsymbol{\Phi}\} \end{bmatrix}\boldsymbol{\Theta} \Rightarrow \boldsymbol{\Theta}=\begin{bmatrix} \mathrm{Re}\{\boldsymbol{\Phi}\} \\ \mathrm{Im}\{\boldsymbol{\Phi}\} \end{bmatrix}^{\#}\begin{bmatrix} \mathrm{Re}\{\boldsymbol{\Psi}\} \\ \mathrm{Im}\{\boldsymbol{\Psi}\} \end{bmatrix}.$$

- Estimate *Markov* parameters: Because $G(z)=\sum_{i=0}^{\infty} Y_i z^{-i}=RQ^{-1}$

$\Rightarrow \left(\sum_{i=0}^{\infty} Y_i z^{-i}\right)\left(\sum_{i=0}^{p} Q_i z^{-i}\right)=\left(\sum_{i=0}^{p} R_i z^{-i}\right)$, compare the coefficients on both sides then $\Rightarrow Y_0=D=R_0,\ Y_k=R_k-\sum_{i=1}^{k} Y_{k-i}Q_i$, $\forall k=1\sim p$, and $Y_k=-\sum_{i=1}^{k} Y_{k-i}Q_i$, $\forall k>p$. ERA method $\Rightarrow$ Balanced State Space form.

- Controllable Canonical-Form realization: $\because y(z_k)=G(z_k)u(z_k)$ $=R(z_k)Q^{-1}(z_k)u(z_k)$. The input of this form can reach each its state. If we define $\tilde{u} \triangleq Q^{-1}u$, where Q is invertible then u completely affects $\tilde{u}$. It is enough that states x are the delay version of $\tilde{u}$. $\because Q\tilde{u}=u \Rightarrow$ $u(z_k)=\left(I_r+Q_1z_k^{-1}+\cdots+Q_pz_k^{-p}\right)\tilde{u}(z_k)$ and $y(z_k)=G(z_k)u(z_k)$ $=R(z_k)Q^{-1}(z_k)u(z_k)=R(z_k)\tilde{u}(z_k)=\left(R_0+R_1z_k^{-1}+\cdots+R_pz_k^{-p}\right)\tilde{u}(z_k)$.

Take $x_1 \triangleq \tilde{u}z^{-p},\ x_2 \triangleq \tilde{u}z^{-p+1},\ \cdots,\ x_p \triangleq \tilde{u}z^{-1}$
$\Rightarrow u(z_k)=\left(I_r+Q_1z_k^{-1}+\cdots+Q_pz_k^{-p}\right)\tilde{u}(z_k)=x_p(z_k)z_k+Q_1x_p(z_k)$ $+Q_2x_{p-1}(z_k)+\cdots+Q_px_1(z_k)$ and $zx_1=x_2,\ zx_2=x_3,\ \cdots,\ zx_{p-1}=x_p$,

$$\Rightarrow z_k \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} 0 & I_r & 0 & \cdots & 0 \\ \vdots & \ddots & I_r & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & I_r \\ -Q_p & -Q_{p-1} & -Q_{p-2} & \cdots & -Q_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ I_r \end{bmatrix} u;$$

$$\Rightarrow z_k x = Ax + Bu(z_k) \Rightarrow y(z_k) = R_0\tilde{u}(z_k) + R_1 x_p(z_k) + R_2 x_{p-1}(z_k) + \cdots$$

$$+R_p x_1(z_k) = R_0 z_k x_p(z_k) + R_1 x_p(z_k) + R_2 x_{p-1}(z_k) + \cdots + R_p x_1(z_k)$$

$$= -R_0\left[Q_1 x_p(z_k) + Q_2 x_{p-1}(z_k) + \cdots + Q_p x_1(z_k)\right] + R_1 x_p(z_k) + R_2 x_{p-1}(z_k)$$

$$+\cdots + R_p x_1(z_k) + R_0 u(z_k) = \underbrace{\begin{bmatrix} R_p - R_0 Q_p & R_{p-1} - R_0 Q_{p-1} & \cdots & R_1 - R_0 Q_1 \end{bmatrix}}_{=C}$$

$$\bullet \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} + \underbrace{R_0}_{=D} u .$$

[Note]: