System Identification : Models

Basic model (for the moment assume SISO)

$$y(t) = \sum_{k=\cancel{0}1}^{\infty} g(k)\,u(t-k) + v(t), \quad t=0,1,\dots$$

A non-standard way to write: Define `q' as the shift operator: $qu(t) = u(t+1)$, $q^{-1}u(t) = u(t-1)$

$$y(t) = \sum_{k=1}^{\infty} g(k)\,u(t-k) = \sum_{k=1}^{\infty} g(k)\left[q^{-k}u(t)\right]$$

$$= \Big[\underbrace{\sum_{k=1}^{\infty} g(k) q^{-k}}_{G(q)}\Big] u(t) = G(q)\,u(t)$$

Model of noise : $v(t)$: We assume

$$v(t) = \sum_{k=0}^{\infty} h(k)\,e(t-k)$$, where $\{e(t)\}$ is white + zero mean

\# Assume $h(0) = 1$.

Hence $v(t) = H(q)\,e(t)$

General Model : $\boxed{y(t) = G(q)u(t) + H(q)e(t)}$

\# We assume $G(q)$ & $H(q)$ are BIBO stable.

\# We assume $G(q)$ & $H(q)$ are rational.

\# Recall if $v(t)$ is stationary (with $\Phi_v(\omega) > 0$) then it has a canonical sp. factorization & $\exists$ $H(q) = 1 + \sum_{k=1}^{\infty} h_k q^{-k}$ s.t. $H(q)$ is stable & min-phase.

Prediction Error Criteria

#Assume $H(q)$ is stably invertible (an automatic consequence if $H(q)$ is obtained from the canonical sp. fact. of $\Phi_v(\omega)$)

→ Then $H^{-1}(q)$ exists & stable.

From the basic model:

$$v(s) = y(s) - G(q)u(s) \qquad s \leq t-1$$

so $v(s)$ are known for $s \leq t-1$

LMS : One step Prediction of $\hat{v}(t)$ given $\{y(1) \cdots y(t-1)\}$:

$$v(t) = H(q)e(t) = \sum_{k=0}^{\infty} h(k) q^{-k} e(t)$$

$$= \sum_{k=0}^{\infty} h(k)e(t-k) = e(t) + \sum_{k=1}^{\infty} h(k)e(t-k)$$

(since $h(0) = 1$)

$$\hat{v}(t/t-1) = 0 + \sum_{k=1}^{\infty} h(k)e(t-k) \quad (\because e_1, \ldots, e_{t-1} \in \mathcal{L}\{y_1 \cdots y_{t-1}\}$$

$\mathcal{L}\{y_1, \ldots, y_{t-1})$

$$= [H(q)-1]e(t) = [H(q)-1]H^{-1}(q)v(t)$$

$$= [1-H^{-1}(q)]v(t)$$

Now if we calculate One-step prediction of $y(t)$: $\hat{y}(t/t-1)$ given $\{y_1, \cdots, y_{t-1}\}$:

$$\hat{y}(t/t-1) = \underbrace{G(q)u(t)}_{\text{known}} + \hat{v}(t/t-1)$$

$$= G(q)\overline{u(t)} + [1-H^{-1}(q)]v(t)$$

$$= G(q)u(t) + [1-H^{-1}(q)][y(t) - G(q)u(t)]$$

$$\boxed{\hat{y}(t/t-1) = H^{-1}(q)G(q)u(t) + [1-H^{-1}(q)]y(t)}$$

Prediction Error (One step) $\equiv$ Innovation

$$y(t) - \hat{y}(t/t-1) = -H^{-1}(q)G(q)u(t) - H^{-1}(q)y(t)$$
$$= e(t)$$

[for $y(t) = G(q)u(t) + H(q)e(t)$
$\Rightarrow e(t) = +H^{-1}(q)y(t)$
$- H^{-1}G(q)u(t)$]

Types of Models

Basic model family:
$$\boxed{\begin{array}{c} y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \\ f_e(x,\theta) \leftarrow \text{pdf of } e(t) \end{array}}$$

Clearly PE also depends on θ:
$$\hat{y}(t/\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + [1 - H^{-1}(q,\theta)]y(t) \qquad ①$$

ARX Model:
$$y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a)$$
$$= b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + e(t)$$

$$\theta = [a_1 \; a_2 \cdots a_{n_a} \; b_1 \cdots b_{n_b}]^T$$

$$G(q,\theta) = \frac{b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}} =: \frac{B(q)}{A(q)}$$

$$H(q,\theta) = \frac{1}{A(q)}$$

ARX identification $\equiv$ Linear Regression or direct calculation:

Putting $G(q,\theta)$ & $H(q,\theta)$ in ①, the PE eqn:
$$\hat{y}(t/\theta) = B(q)u(t) + [1 - A(q)]y(t)$$

Define:
$$\phi(t) = [-y(t-1) \cdots -y(t-n_a) \; u(t-1) \cdots u(t-n_b)]^T$$

$$\hat{y}(t/\theta) = \theta^T \phi(t) = \phi^T(t)\theta \qquad t = 0, 1 \ldots$$

(we can use std. techniques)

ARMAX Model

$$y(t)+a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \cdots$$
$$\cdots + b_{n_b} u(t-n_b) + e(t) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c)$$

$$A(q)y(t) = B(q)u(t) + C(q)e(t)$$
$$\theta = [a_1 \cdots a_{n_a}\ b_1 \cdots b_{n_b}\ c_1 \cdots c_{n_c}]^T$$

Predictor Expression: $$\hat{y}(t/\theta) = \frac{B(q)}{C(q)}\,u(t) + \left[1 - \frac{A(q)}{C(q)}\right] y(t) \quad \textcircled{1}$$

Note: It might be hard to start at $t=0$ [knowledge of $\{\hat{y}(0/\theta) \cdots \hat{y}(-n_c+1/\theta)$, $y(0) \cdots y(-n^*+1), u(0), \ldots, u(-n_b+1)\}$ $n^* = \max(n_c, n_a)$ is req]

Defining $\varepsilon(t,\theta) = y(t) - \hat{y}(t/\theta)$

& $\phi(t,\theta) = [-y(t-1) \cdots -y(t-n_a)\ u(t-1) \cdots u(t-n_b)\ \ \varepsilon(t-1,\theta) \cdots \varepsilon(t-n_c,\theta)]^T$.

$$\textcircled{1} \underset{\substack{\uparrow\\ \text{Exercise}}}{\Longleftrightarrow} \underline{\hat{y}(t/\theta) = \phi(t,\theta)\,\theta}$$

Pseudo-linear Regression since ϕ is a $f^{\underline{n}}$ of θ.

Many such models: most can be written in linear/pseudo-linear form.

State Space Models

$$x(t+1) = A(\theta)\,x(t) + B(\theta)\,u(t) + w(t)$$
$$y(t) = C(\theta)\,x(t) + v(t)$$

$\Updownarrow$ we would like to write

$$y(t) = G(q,\theta)\,u(t) + \boxed{\ ?\ }\,e(t) \quad ②$$
where $G(q,\theta) = C(\theta)\,[qI - A(\theta)]^{-1}B(\theta)$

We can use our knowledge of Kalman filter / innovations representation to directly write:

$y(t)$ in term of innovations:

$$\hat{x}(t+1/\theta) = A(\theta)\,\hat{x}(t,\theta) + B(\theta)\,u(t) + K(\theta)\,e(t)$$
$$y(t) = C(\theta)\,\hat{x}(t,\theta) + e(t)$$

$\rightarrow$ Now we can fill the blank in ② :

$$\boxed{y(t) = G(q,\theta)\,u(t) + H(q,\theta)\,e(t)}$$

where $G(q,\theta) = C(\theta)\,[qI - A(\theta)]^{-1}B(\theta)$

$H(q,\theta) = C(\theta)\,[qI - A(\theta)]^{-1}K(\theta) + I$

Similarly,
$$\boxed{\begin{aligned}\hat{y}(t/\theta) &= C(\theta)\,[qI - A(\theta) + K(\theta)C(\theta)]^{-1}B(\theta)\,u(t)\\ &\quad + C(\theta)\,[qI - A(\theta) + K(\theta)C(\theta)]^{-1}K(\theta)\,y(t)\end{aligned}}$$

where, $K(\theta)$ is computed using the (by now familiar!) Discrete time Riccati Equation.

Clearly, a parameterization in term of ARE might be hard to estimate.

$\rightarrow$ Sometimes easier to directly estimate $K(\theta)$.

Companion Form Parameterization / Observer Canonical form

set $\theta^T = [a_1\ a_2\ a_3\ \ b_1\ b_2\ b_3\ \ k_1\ k_2\ k_3]$

& $A(\theta) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix}$ $\quad B(\theta) = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ $\quad K(\theta) = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}$

$C(\theta) = [1\ 0\ 0]$

Then $C(\theta)[qI - A(\theta)]^{-1}B(\theta) = \dfrac{b_1q^{-1} + b_2q^{-2} + b_3q^{-3}}{1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3}}$

$$C(\theta)[qI - A(\theta)]^{-1}K(\theta) = \frac{k_1q^{-1} + k_2q^{-2} + k_3q^{-3}}{1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3}}$$

$$1 + C(\theta)[qI - A(\theta)]^{-1}K(\theta) = \frac{1 + c_1q^{-1} + c_2q^{-2} + c_3q^{-3}}{1 + a_1q^{-1} \cdots\cdots}$$

where $c_i^0 = a_i^0 + k_i^0$ $\quad i = 1, 2, 3.$

$\equiv$ ARMAX model.

Note: Unlike the companion form above, S.S. models may arise from physical laws $\Rightarrow$ might lead to lesser parameters.

Estimation of State Space Models

$$(*)\begin{cases} x(t+1) = Ax(t) + Bu(t) + w(t) \\ y(t) = Cx(t) + Du(t) + v(t) \end{cases}$$

Assume ω, v white noises.

Note: $y(t)$ & $u(t)$ are observed. But there are infinitely many correct realizations that can link them.

Somehow we need to fix the coordinate basis of x.

Temporarily assume x is also measured along with u & y.

Then $Y(t) = \begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix}$ $\quad \Theta = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$

$\Phi(t) = \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$ $\quad E(t) = \begin{bmatrix} w(t) \\ v(t) \end{bmatrix}$

Then $\circledast$ is equivalent to

$$Y(t) = \Theta \Phi(t) + E(t)$$

→ Estimate θ using least-squares.

⇝ $\mathrm{Cov}(E(t))$ can be also estimated from calculating $[Y(t) - \Theta \Phi(t)]$ for sample t's. → will give us $\left\langle \begin{bmatrix} w \\ v \end{bmatrix}, \begin{bmatrix} w \\ v \end{bmatrix} \right\rangle$.

Q. In reality how to compute x ?

Realization from Impulse Response (HO-Kalman Algo)

$x(t+1) = Ax(t) + Bu(t)$ $\quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p$

$y(t) = Cx(t) + Du(t)$ $\quad t = 0, 1, \cdots$

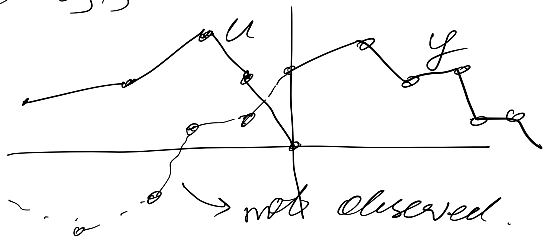
We assume (A, B, C) minimal. (Reachable + Observable)

$G(z) = D + C(zI-A)^{-1}B$; $G_t = \begin{cases} D & t=0 \\ CA^{t-1}B & t=1,2,\cdots \end{cases}$

Problem: Given G_t, find dimension n & (A,B,C,D) upto similarity transforms.

#Assume $u(i) = \begin{cases} u(i) & i<0 \\ 0 & i \geq 0 \end{cases}$ We observe $y(t)$ from $t=0,1,\cdots$

Then $y(t) = \sum_{i=-\infty}^{-1} G_{t-i}\, u(i)$, $t=0,1,\cdots$



Define the block Hankel matrices:

$$H := \begin{bmatrix} G_1 & G_2 & G_3 & \cdots \\ G_2 & G_3 & G_4 & \cdots \\ G_3 & G_4 & G_5 & \cdots \\ \vdots & & & \end{bmatrix}_{\text{infinite dim.}} \qquad \bar{H} = \begin{bmatrix} G_2 & G_3 & \cdots \\ G_3 & G_4 & \cdots \\ G_4 & G_5 & \cdots \\ \vdots & & \end{bmatrix}$$

Also, $y_+ := \begin{bmatrix} y(0) \\ y(1) \\ \vdots \end{bmatrix}$ $u_- = \begin{bmatrix} u(-1) \\ u(-2) \\ \vdots \end{bmatrix}$, then $\boxed{\underset{\text{future output}}{y_+} = H \underset{\text{past input}}{u_-}}$

Define $\mathcal{C} = [B \ AB \ A^2B \ \cdots]$ $\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \end{bmatrix}$

Thm: Let $\{A,B,C\}$ be minimal. Then the following hold:

(i) $\operatorname{rank}(H) = n$.

(ii) $H = \mathcal{O}\mathcal{C}$ $(= \mathcal{O}TT^{-1}\mathcal{C}$ for non-singular $T.)$
& $\bar{H} = \mathcal{O}A\mathcal{C}$

(ii) H is shift invariant i.e

$$H^{\uparrow} = \mathcal{O}^{\uparrow}\mathcal{C} = \mathcal{O}A.\,\mathcal{C} = \mathcal{O}.\,A\mathcal{C} = \mathcal{O}\mathcal{C}^{\leftarrow} = H^{\leftarrow}$$

where $(\cdot)^{\uparrow}$ is upward shift, $(\cdot)^{\leftarrow}$ is left shift.

Note: If $x(0) = \mathcal{C}u_-$, then $y_+ = \mathcal{O}x(0)$ (from $y_+ = Hu_-$)

Define $H_{k,l} = \begin{bmatrix} G_1 & \cdots & G_l \\ \vdots & & \vdots \\ G_k & \cdots & G_{k+l-1} \end{bmatrix} \in \mathbb{R}^{kp \times lm}$; $\bar{H}_{k,l} = \begin{bmatrix} G_2 & \cdots & G_{l+1} \\ \vdots & & \\ G_{k+1} & \cdots & G_{k+l} \end{bmatrix}$

$\mathcal{O}_k = \begin{bmatrix} C \\ \vdots \\ CA^{k-1} \end{bmatrix}$ $\quad \mathcal{C}_l = [B \;\; AB \cdots A^{l-1}B]$ $\quad$ where $\underline{k, l > n}$.

Usually, $n < k < l$.

Note: 'n' is not known. So just guess a "large enough" k, l.

$\Rightarrow$ rank $(\mathcal{O}_k)$ = rank $\mathcal{C}_l = n$.

We use, $H_{k,l} = \mathcal{O}_k\mathcal{C}_l$ & $\bar{H} = \mathcal{O}_k A \mathcal{C}_l$

$$\mathcal{O}_k = \begin{bmatrix} C \\ \mathcal{O}_{k-1}A \end{bmatrix} \quad \text{and} \quad \mathcal{C}_l = [B \;\; A\mathcal{C}_{l-1}]$$

First we get the rank n-estimate of H

$$\hat{H}_n = \underset{\text{rank}(\hat{H}=n)}{\arg\min} \|\hat{H}_n - H_{k,l}\|_2 = U_n \Sigma_n V_n^T \leftarrow \underline{n \times n}$$

If $H_{k,l} = [U_n \;\; U_s]\begin{bmatrix} \Sigma_n & \\ & \Sigma_s \end{bmatrix}\begin{bmatrix} V_n^T \\ V_s^T \end{bmatrix}$

(If no noise, $\Sigma_s = 0$, otherwise n can be estimated from the range of singular values.

Then estimates of $\mathcal{O}_k$ & $\mathcal{C}_l$ can be calculated

from $\hat{O}_k = U_n \Sigma_n^{1/2}$ & $\hat{C}_l = \Sigma_n^{1/2} V_n^T$

Finally, A can be estimated as

$$\hat{A} = (\hat{O}_k)^{\dagger} \bar{H} (\hat{C}_l)^{\dagger} = \Sigma_n^{-1/2} U_n^T \bar{H} V_n \Sigma_n^{-1/2} \quad \text{(Exercise)}$$

B, C can be estimated easily:

$$\hat{B} = \hat{C}_l(:, 1:m) \qquad \hat{C} = \hat{O}_k(1:p, :)$$

Example: Suppose $G_t = G_{t-1} + G_{t-2}$ ∧ $t = 2, 3, \ldots$ $G_0 = 0, G_1 = 1$

The tr. fn: $G(z) = \dfrac{z}{z^2 - z - 1}$ (This generate the Fibonacci seq. for impulse input)

Directly we can calculate: $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$C = [1 \;\; 0]$, $D = 0$.

Now, using our algo: Take $k, l = 5$.

$$H_{5,5} = \left[\begin{array}{ccccc} G_1 & G_2 & \cdots & & \\ 1 & 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & 5 & 8 \\ 2 & 3 & 5 & 8 & 13 \\ 3 & 5 & 8 & 13 & 21 \\ 5 & 8 & 13 & 21 & 34 \end{array}\right] = U \left[\begin{array}{cc|ccc} 54.56 & & & & \\ & 0.439 & & 0 & \\ \hline & & 2\times10^{-15} & & \\ & 0 & & 8\times10^{-16} & \\ & & & & 6\times10^{-18} \end{array}\right] V^T$$

So we take $n = 2$, $\hat{A} = \begin{bmatrix} 1.62 & 0.02 \\ 0.02 & -0.62 \end{bmatrix}$ $\hat{B} = \begin{bmatrix} 0.85 \\ -0.52 \end{bmatrix}$

$\hat{C} = [0.85 \;\; -0.52]$ $D = 0$.

Check $\hat{C}(zI - \hat{A})^{-1}\hat{B} = \dfrac{z}{z^2 - z - 1}$

:

Q. How to estimate n, & $\{A, B, C, D\}$ from I/P, O/P data?

$\{A, B, C, D\}$ from data matrices - Deterministic Case

Assume : $\left.\begin{cases} u(0) & u(1) \cdots\ u(i+\ -2) \\ y(0) & y(1) \cdots\ y(i+\ -2) \end{cases}\right\}$ given

Define : $$U_{0|\ -1} = \begin{bmatrix} u(0) & u(1) & \cdots & u(\ -1) \\ u(1) & u(2) & \cdots & u(\ ) \\ \vdots & & & \\ u(i-1) & u(k) & \cdots & u(i+\ -2) \end{bmatrix} \in \mathbb{R}^{km\times N}$$

$$Y_{0|\ -1} = \begin{bmatrix} y(0) & \cdots & y(\ -1) \\ \vdots & & \\ y(k-1) & \cdots & y(\ +\ -2) \end{bmatrix} \in l$$

Then $$\underbrace{\begin{bmatrix} y(t) \\ y(t+1) \\ \vdots \\ y(t+\ -1) \end{bmatrix}}_{Y\ (t)} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ '} \end{bmatrix} x(t) + \underbrace{\begin{bmatrix} D & & 0 \\ CB & D & \\ & \ddots & \ddots \\ CA^{\ 2}B & CB & D \end{bmatrix}}_{\Psi} \underbrace{\begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+\ -1) \end{bmatrix}}_{U\ (t)}$$

So $\quad Y\ (t) = \mathcal{O}\, x(t) + \Psi\ U\ (t)$

Also, $\left.\begin{aligned} U_{0|\ -1} &= [u(0)\ \ u(1)\ \cdots\ u(\ -1)] \\ Y_{0|\ -1} &= [y(0)\ \ y(1)\ \cdots\ y(\ -1)] \end{aligned}\right\}$ past I/Ps & O/Ps

So, $$\boxed{Y_{0|\ -1} = \mathcal{O}\, X_0 + \Psi\ U_{0|\ -1}} \text{ — } ①$$

where $\quad X_0 = [x(0)\ \ x(1)\ \cdots\ x(\ -1)]_{n\times N}$

$$\overset{\text{Future I/Ps \& O/Ps}}{}$$

Similarly, define : $\begin{cases} U_{\ |2\ -1} = [u(\ )\ \ u(\ +1) \cdots u(\ \ -1)] \\ Y_{\ |2\ -1} = [y(k)\ \ y(\ +1) \cdots y_k(\ \ -1)] \end{cases}$

Then, $\boxed{Y_{k|2k-1} = O\, X + \Psi\, U_{|\ -1}}$ — ②

where $X_k = [x(k)\ \ x(k+1) \cdots x(k+N-1)]_{n\times N}$

Basic Idea : 1) Calculate $O_k x_k$ from ① & ②

2) Factorize $O_k x_k$ to calculate x_k using SVD.

Outline of Step 1:

Shorthand: : $W_p = \begin{bmatrix} U_{0|k-1} \\ Y_{0|k-1} \end{bmatrix} =: \begin{bmatrix} U_p \\ Y_p \end{bmatrix}$; $\begin{bmatrix} U_f \\ Y_f \end{bmatrix} := \begin{bmatrix} U_{k|2k-1} \\ Y_{k|2k-1} \end{bmatrix}$

Now do LQ decomposition of :

$$\begin{bmatrix} U_f \\ W_p \\ Y_f \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{bmatrix} \text{ — } (*)$$

Using (*), we can derive (some work req.)

$$Y_f = \left(R_{31} - R_{32} R_{22}^{\dagger} R_{21}\right) R_{11}^{-1} U_f + R_{32} R_{22}^{\dagger} W_p \quad ③$$

Compare with ②:

$$Y_f = \Psi_k U_f + O_k x_k$$

Using careful rank considerations
(lot of work req.)

Lemma (without proof): 1) $O_k x_k = R_{32} R_{22}^{\dagger} W_p$

2) $\Psi_k = \left(R_{31} - R_{32} R_{22}^{\dagger} R_{21}\right) R_{11}^{-1}$

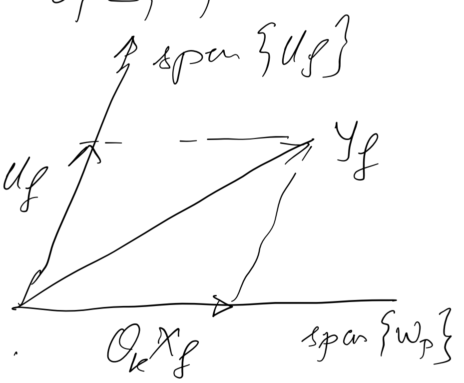
From this eqn (1), X_k can be calculated by:

1) Define $\xi = R_{32} R_{22}^{\dagger} W_p \; (= O_k X_k)$

2) $\xi = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_1 V_1^T$

3) $O_k = U_1 \Sigma_1^{1/2}$

4) $X_k = \Sigma_1^{1/2} V_1^T$



Computation of $\{A, B, C, D\}$:

Define $\bar{X}_{k+1} := [x(k+1) \cdots x(k+N-1)]$

$\bar{X}_k := [x(k) \cdots x(k+N-2)]$

$\bar{U}_{k|k} := [u(k) \cdots u(k+N-2)]$

$\bar{Y}_{k|k} := [y(k) \cdots y(k+N-2)]$

Then
$$\begin{bmatrix} \bar{X}_{k+1} \\ \bar{Y}_{k|k} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \bar{X}_k \\ \bar{U}_{k|k} \end{bmatrix}$$

$$\begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} = \left(\begin{bmatrix} \bar{X}_{k+1} \\ \bar{Y}_{k|k} \end{bmatrix} \begin{bmatrix} \bar{X}_k \\ \bar{U}_{k|k} \end{bmatrix}^T \right) \left(\begin{bmatrix} \bar{X}_k \\ \bar{U}_{k|k} \end{bmatrix} \begin{bmatrix} \bar{X}_k \\ \bar{U}_{k|k} \end{bmatrix}^T \right)^{-1}$$

To show $O_k x_k \in \text{span}\{W_p\}$,

$\longrightarrow$ exercise

$$X_k = A^k X_0 + \bar{C}_k U_{0|k-1} \quad \left(\text{where } \bar{C}_k = [A^{k-1}B \ \ A^{k-2}B \ \dots \ B]\right)$$

Therefore, $X_0 = O_k^{\dagger} Y_{0|k-1} - O_k^{\dagger} \Psi_k U_{0|k-1}$ (from ①)

So $X_k = A^k\left[O_k^{\dagger} Y_{0|k-1} - O_k^{\dagger} \Psi_k U_{0|k-1}\right] + \bar{C}_k U_{0|k-1}$

$\longrightarrow$ ③

$\Rightarrow X_k \in \text{span}\{W_p\}$

Not here

Also, $\Rightarrow \mathcal{O}_{k-1} A = \mathcal{O}_k(p+1:kp, :)$

To get unique LS. solⁿ. for A, $\mathcal{O}_{k-1}$ must be full col. rank., so $p(k-1) \geq n$.
For single input, $p=1$, $\boxed{k > n}$

Similarly, $\Leftrightarrow A\mathcal{C}_{l-1} = \mathcal{C}_l(:, m+1: lm)$
where we require $m(l-1) \geq n$ & for $m=1$, $\boxed{l > n}$

▷ Derive a k-step ahead predictor for y_t from past $\{y_t\}$ & $\{u_t\}$ data

In general: $y(t) = \sum_{j=0}^{\infty} \left[h_u(j)u(t-j) + h_e(j)e(t-j)\right]$

For the k-step ahead predictio

$$\hat{y}(t+k-1 / t-1) = \sum_{j=1}^{\infty} h_u(j)\, u(t-j) + h_e(j)\, e(t-j)$$

$\hookrightarrow \circledast$

$\hookrightarrow$ Note: We don't try to predict $u(j)$; $j=t \dots, t+k+1$

Basic Idea:

So in principle, X_f can be — from ③ if we know the $\{A, B_{\_}, C_{\_}\}$.

Putting ③ in ②:

$$Y_{k/2k-1} = \mathcal{O}_k A^k \mathcal{O}_k^+ Y_{0/k-1} - \mathcal{O}_k A^k \mathcal{O}_k^+ \Psi_k U_{0/k-1} + \mathcal{O}_k \bar{\Phi}_k U_{0/k-1} + \Psi_k U_{k/2k-1}$$

$$= F Y_{0/k-1} + H U_{0/k-1} + \Psi_k U_{k/2k-1}$$

ROUGHLY: some estimate of F, H, Ψ_k can be computed to get X_k using ③

#For practical reasons, we curtail this: with finite past data points: (S_1, S_2)

So $\textcircled{\wp}$ simplifies to:

$$\hat{y}(t+k-1/t-1) = \alpha_1 y(t-1) + \cdots + \alpha_{S_1} y(t-S_1) + \beta_1 u(t-1) + \cdots + \beta_{S_2} u(t-S_2)$$

$$= \theta_k^T \phi_S(t)$$

where $\theta_k^T = [\alpha_1 \cdots \alpha_{S_1} \;\; \beta_1 \cdots \beta_{S_2}]$

$$\phi_S(t) = [y_1(t-1) \cdots y(t-S_1) \;\; u(t-1) \cdots u(t-S_2)]^T$$

$$\text{Define: } \hat{y}_r(t) = \begin{array}{r} \\ \\ \\ \textcircled{\wp}\longrightarrow \end{array}\begin{bmatrix} \hat{y}(t/t-1) \\ \vdots \\ \\ \hat{y}(t+r-1/t-1) \end{bmatrix} = \begin{bmatrix} \theta_1^T \phi_S(t) \\ , \\ \vdots \\ \theta_r^T \phi_S(t) \end{bmatrix}$$

$\hookrightarrow$ same $\phi_S(t)$ since thats all the data we have.

i.e. $\hat{y}_r(t) = \Theta \phi_S(t)$

where $\Theta = [\theta_1 \cdots \theta_r]^T$

$$\& \quad \hat{Y} = \begin{bmatrix} \hat{y}(1/0) & \hat{y}(2/1) & & \hat{y}(N/N-1) \\ \hat{y}(2/0) & \hat{y}(3/1) & \cdots & \hat{y}(N+1/N-1) \\ \vdots & \vdots & & \vdots \\ \hat{y}(r/0) & \hat{y}(1+r/1) & & \hat{y}(N+r-1/N-1) \end{bmatrix}$$

System Identification - Deterministic Case

Problem: Given s measurements of the input $u_k \in \mathbb{R}^m$ and output $y_k \in \mathbb{R}^l$ generated by unknown deterministic system of order 'n':

$$x_{k+1} = Ax_k + Bu_k \quad \text{——} ①$$

$$y_k = Cx_k + Du_k \quad \text{——} ②$$

Determine : 1) the order n

2) A, B, C, D (upto sim. tr.)

From ②,

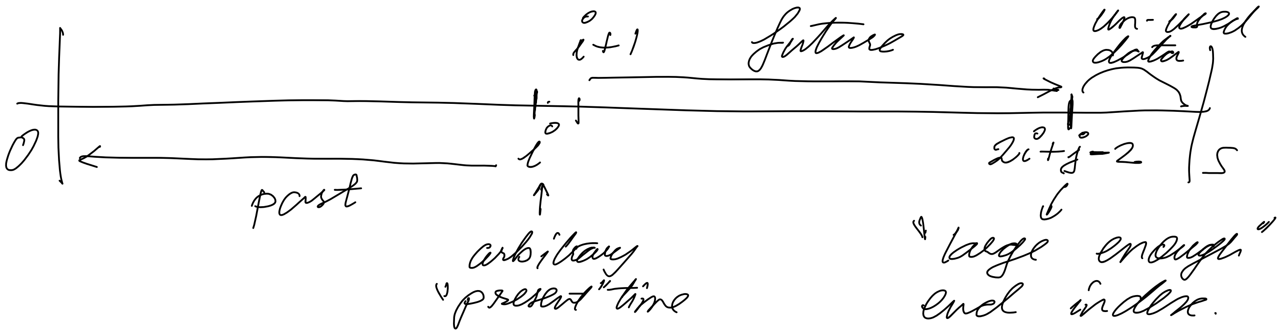
$$\begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+q} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} x_k + \begin{bmatrix} D & & 0 \\ CB & D & \\ \vdots & \ddots & \ddots \\ CA^{q-2}B & CB & D \end{bmatrix} \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+q} \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} y_k & y_{k+1} & \cdots & y_{k+j-1} \\ \vdots & \vdots & & \vdots \\ y_{k+q-1} & y_{k+q} & & y_{k+q+j-2} \end{bmatrix}}_{Y_{k/q}} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix}}_{\Gamma_q} \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ x_k & x_{k+1} & \cdots & x_{k+j-1} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{X_k}$$

$$+ \underbrace{\begin{bmatrix} D & & 0 \\ CB & D & \\ & \ddots & \ddots \\ CA^{q-2}B & CB & D \end{bmatrix}}_{H_q} \underbrace{\begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+j-1} \\ \vdots & \vdots & & \vdots \\ u_{k+q-1} & u_{k+q} & & u_{k+q+j-2} \end{bmatrix}}_{U_{k/q}}$$

$\Leftrightarrow$ $\boxed{Y_{k/q} = \Gamma_q X_k + H_q U_{k/q}}$ $\longrightarrow$ Ⓓ

#We assume data collection starts from k = 0 and ends at s.



#We define the above matrices on the two seperate data blocks :

1> **Past** : $k=0,\ q=i$: $U_p := U_{0/i-1}$

$Y_p := Y_{0/i-1}$

$$X_p := X_0 = [x_0\ x_1\ \cdots\ x_{j-1}]$$

$$\boxed{Y_p = \Gamma_i X_p + H_i U_p} \longrightarrow ①$$

2> **Future** : $k=i,\ q=2i-1$: $U_f := U_{i/2i-1}$

$Y_f := Y_{i/2i-1}$

$$X_f := X_i = [x_i\ x_{i+1}\ \cdots\ x_{i+j-1}]$$

$$\boxed{Y_f = \Gamma_i X_f + H_i U_f} \longrightarrow ②$$

stacked state eqn:

$$x_{k+1} = A x_k + B u_k$$

$$x_{k+2} = A[A x_k + B u_k] + B u_{k+1}$$

$$= A^2 x_k + B u_{k+1} + AB u_k$$

$$x_{k+q} = A^q x_k + \underbrace{[A^{q-1}B \cdots\ AB\ B]}_{\Delta_q} \begin{bmatrix} u_k \\ \vdots \\ u_{k+q} \end{bmatrix}$$

Stacking these eqns: for k = i to i+j-1

$$k=0,\ q=i:\quad x_i = A^i x_0 + \Delta_i \begin{bmatrix} u_0 \\ u_i \end{bmatrix}$$

$$k=1,\ q=i+1:\quad x_{i+1} = A^i x_1 + \Delta_i \begin{bmatrix} u_1 \\ \vdots \\ u_{1+i} \end{bmatrix}$$

$$\vdots$$

$$k=j-1,\ q=i+j-1:\quad x_{i+j-1} = A^i x_{j-1} + \Delta_i \begin{bmatrix} u_{j-1} \\ \vdots \\ u_{i+j-1} \end{bmatrix}$$

$$\begin{bmatrix} x_i & x_{i+1} & \cdots & x_{i+j-1} \end{bmatrix} = A^i \begin{bmatrix} x_0 & x_1 & \cdots & x_{i+j-1} \end{bmatrix}$$

$$+ \Delta_i \begin{bmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ \vdots & \vdots & & \vdots \\ u_i & u_{i+1} & & u_{i+j-1} \end{bmatrix}$$

$$\Leftrightarrow \quad X_i = A^i X_0 + \Delta_i U_{0|i}$$

$$\Leftrightarrow \boxed{X_f = A^i X_p + \Delta_i U_p} \text{ --- } ③$$

We can eliminate X_p & X_f from ①,②,③

$$X_f = A^i X_p + \Delta_i U_p$$

$$= A^i \left[\Gamma_i^+ Y_p - \Gamma_i^+ H_i U_p \right] + \Delta_i U_p$$

$$(\text{from } Y_p = \Gamma_i X_p + H_i U_p)$$

$$= \left[\Delta_i - A^i \Gamma_i^{+} H_i\right] U_p + \left[A^i \Gamma_i^{+}\right] Y_p \longrightarrow \circledast$$

Using $\circledast$ in ②,

$$Y_f = \Gamma_i \left[\Delta_i - A^i \Gamma_i^{+} H_i\right] U_p + \Gamma_i A^i \Gamma_i^{+} Y_p + H_i U_f$$

Notatn: $W_p := \begin{bmatrix} U_p \\ Y_p \end{bmatrix}$ $L_p = \left[\Delta_i - A^i \Gamma_i^{+} H_i \,\middle|\, A^i \Gamma_i^{+}\right]$

$$\Rightarrow \quad Y_f = \underbrace{\Gamma_i L_p W_p}_{\circledast} + H_i U_f \quad \Big| \quad X_f := L_p W_p$$

Now compare with ② again:

$$Y_f = \Gamma_i X_f + H_i U_f$$

#If we can calculate $\circledast$, we know $\Gamma_i X_f$.
We can factorize arbitrarily to create X_f & Γ_i.

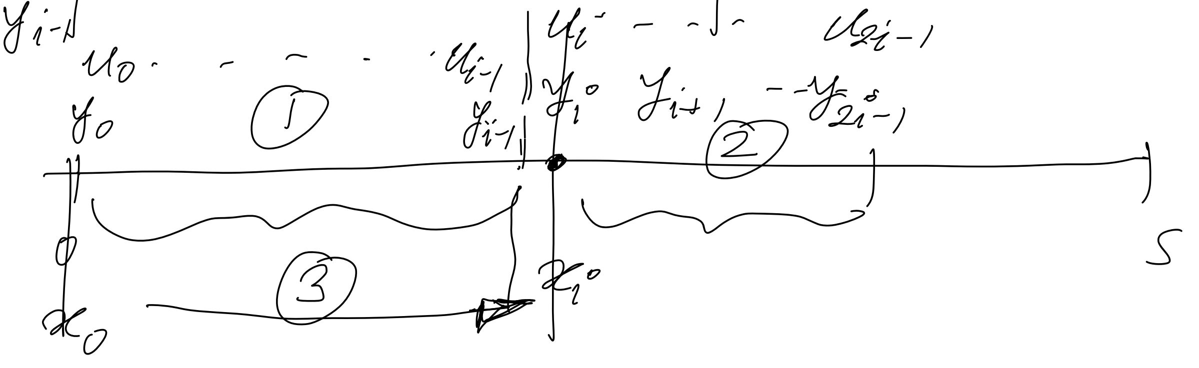
Q. How to calculate $\circledast$ from data?
— Ans: Clever Projections.

Equivalent Unstacked Eqns:

$$\begin{bmatrix} y_i \\ y_{i+1} \\ \vdots \\ y_{2i-1} \end{bmatrix} = \Gamma_i x_i + H_i \begin{bmatrix} u_i \\ \vdots \\ u_{2i-1} \end{bmatrix}$$

$\overline{Y}_f = \Gamma_i^o x_i^o + H_i^o \overline{U}_f$ — ② $\quad x_i^o = A^i x_0 + \Delta_i^o \overline{U}_p$ ↳ ③

$\begin{bmatrix} y_0 \\ \vdots \\ y_{i-1} \end{bmatrix} = \overline{Y}_p = \Gamma_i x_0 + H_i^o \overline{U}_p$ — ①



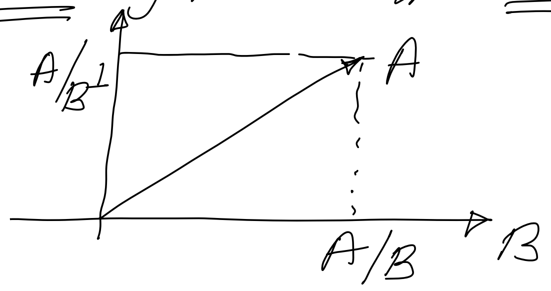
Review of Projection:

Orthogonal Projection: $A \in \mathbb{R}^{p \times j}$, $B \in \mathbb{R}^{q \times j}$ & $C \in \mathbb{R}^{r \times j}$

$\# \;\; A/B := A\Pi_B = AB^T(BB^T)^{\dagger}B =: L_B B$

[Note: Projection of rows of A onto the row space of B]

$j = 2$



$\# \;\; A/B^{\perp} = A\Pi_{B^{\perp}} = I_i^o - \Pi_B =: L_{B^{\perp}} B^{\perp}$

$\# \;\; A = A\Pi_B + A\Pi_{B^{\perp}} = L_B B + L_{B^{\perp}} B^{\perp}$

Using RQ decomp: (Q^T: orthonormal, R: lower triangular)

Let $A = R_A Q^T$; $B = R_B Q^T$

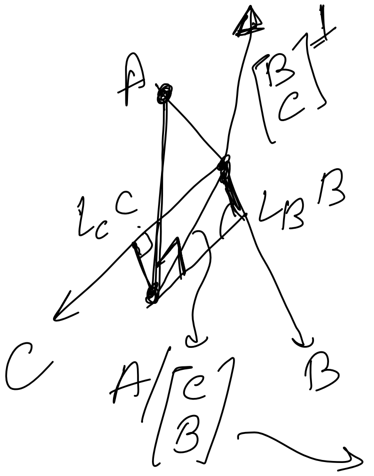
$$\begin{bmatrix} A \\ B \end{bmatrix} = \left[\begin{array}{c|c} R_A & 0 \\ \hline \multicolumn{2}{c}{R_B} \end{array}\right] \begin{bmatrix} \quad Q^T \quad \end{bmatrix}$$

$$A/B = \left[R_A Q^T Q R_B^T\right]\left[R_B Q^T Q R_B^T\right]^{+} R_B Q^T$$

$$= R_A R_B^T \left[R_B R_B^T\right]^{+} R_B Q^T$$

(Oblique) Projection: Project the rows of A onto the rows of B and C

In general: $A = L_B B + L_C C + L_{B^\perp C^\perp}\begin{bmatrix} B \\ C \end{bmatrix}^\perp$



Projection of row space of A along row sp. of B onto row sp. of C.

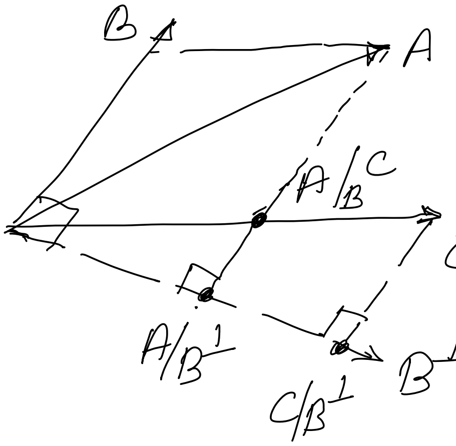
$= L_C C =: A/_B C$

$$A/\begin{bmatrix} C \\ B \end{bmatrix} = A\begin{bmatrix} C^T & B^T \end{bmatrix}\begin{bmatrix} CC^T & CB^T \\ BC^T & -BB^T \end{bmatrix}^{+}\begin{bmatrix} C \\ B \end{bmatrix}$$

$$\left.\begin{array}{l} = \left[A\begin{bmatrix} C^T & B^T \end{bmatrix}\begin{bmatrix} CC^T & CB^T \\ BC^T & BB^T \end{bmatrix}^{+}\right]_{\text{first } r \text{ columns}} \cdot C \\ + \left[\qquad '' \qquad \right]_{\text{last } q \text{ columns}} \cdot B \end{array}\right\} \begin{array}{l} = A/_B C \\ = A/_C B \end{array}$$

Note: $B/_B C = 0$; $C/_B C = C$

FACT: $A/_B C = [A/B^\perp][C/B^\perp]^+ C$



Clearly:

$A/_B C = [M] C$ — ①

$A/B^\perp = [M] C/B^\perp$ — ②

From ②,

$[M] = A/B^\perp [C/B^\perp]^+$

$\Rightarrow A/_B C = A/B^\perp [C/B^\perp]^+ C$ — (*)

Using QR: $\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} R_A \\ R_B \\ R_C \end{bmatrix} \begin{bmatrix} Q^T \end{bmatrix}$

$A/B^\perp = R_A [I - R_B^T (R_B R_B^T]^+ R_B] Q^T$

$C/B^\perp = R_C [I - R_B^T [R_B R_B^T)^+ R_B] Q^T$

One could use these expressions in (*).

Continue with Sys ID

Recall the final eqn: $Y_f = \Gamma_i^\circ L_p W_p + H_i^\circ U_f$

and $X_f = L_p W_p$ — (*)

Claim: $\xi := Y_f/_{U_f} W_p = \Gamma_i^\circ L_p W_p = \Gamma_i^\circ X_f$

Proof: From (*), $Y_f \Pi_{U_f^\perp} = \Gamma_i^\circ L_p W_p \Pi_{U_f^\perp} + H_i^\circ \underbrace{U_f \Pi_{U_f^\perp}}_{=0}$

or $\mathcal{Y}_f/\mathcal{U}_f^{\perp} = \Gamma_i^{o} L_p W_p/\mathcal{U}_f^{\perp}$

Post multiply both sides by $[W_p/\mathcal{U}_f^{\perp}]^{\dagger} \cdot W_p$

$$\underbrace{[\mathcal{Y}_f/\mathcal{U}_f^{\perp}][W_p/\mathcal{U}_f^{\perp}]^{\dagger} W_p}_{\|} = \Gamma_i L_p \underbrace{[W_p/\mathcal{U}_f^{\perp}][W_p/\mathcal{U}_f^{\perp}]^{\dagger} W_p}_{W_p}$$

or $\mathcal{Y}_f/_{\mathcal{U}_f} W_p = \Gamma_i^{o} L_p W_p$

Q. Why is $[W_p/\mathcal{U}_f^{\perp}][W_p/\mathcal{U}_f^{\perp}]^{\dagger} W_p = W_p$?

→ Answer postponed. ←

For the moment assume "claim" is true.

Then, let

$$\xi = [\mathcal{U}_1 \ \ \mathcal{U}_2]\begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

Linked

$$= \mathcal{U}_1 S_1 V_1^T$$

$\#$ n = size of S_1. [Proof postponed] ←

$\#$ $\Gamma_i^{o} X_f = \xi = \underbrace{\mathcal{U}_1 S_1^{1/2} T}_{\Gamma_i} \underbrace{T^{-1} S_1^{1/2} V_1^T}_{X_f}$

$\Rightarrow \Gamma_i^{o} = \mathcal{U}_1 S_1^{1/2} T$. and

$X_f = T^{-1} S_1^{1/2} V_1^T$

Note: 1> rank$(\mathcal{Y}_f/_{\mathcal{U}_f} W_p) = n$

2> row sp$(\mathcal{Y}_f/_{\mathcal{U}_f} W_p)$ = row sp(X_f)

3> col. sp$(\mathcal{Y}_f/_{\mathcal{U}_f} W_p)$ = col sp(Γ_i^{o})

4) row sp X_f = row sp. $\begin{bmatrix} U_p \\ Y_p \end{bmatrix} \cap$ row sp. $\begin{bmatrix} U_f \\ Y_f \end{bmatrix}$

$$\left[\; X_f = \Gamma_i^{+} Y_f /_{U_f} W_p \; . \quad \text{Also,} \quad X_f = \Gamma_i^{+} \underbrace{Y_f}_{\uparrow} - \Gamma_i^{+} H_i \underbrace{U_f}_{\uparrow} \;\right]$$

Computing A, B, C, D

Algo 1: Using the states

Recall:

$$\begin{array}{l} i \times j \in U_p \left\{ \begin{array}{l} \\ \\ \end{array} \right. \\ \\ i \times j \in U_f \left\{ \begin{array}{l} \\ \\ \end{array} \right. \end{array} \begin{bmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ \vdots & & & \\ u_{i-1} & u_i & \cdots & u_{i+j-2} \\ \hline u_i & u_{i+1} & \cdots & u_{i+j-1} \\ \hline \vdots & & & \\ u_{2i-1} & u_{2i} & \cdots & u_{2i+j-2} \end{bmatrix} \begin{array}{l} \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} U_p^{+} \in (i+1) \times j \\ \left. \begin{array}{l} \\ \end{array} \right\} U_f^{-} \in (i-1) \times j \end{array}$$

Similarly define: X_p^{+}, Y_f^{-}. Then similar arguments as above,

$$\underline{Eq} := Y_f^{-} /_{U_f^{-}} W_p^{+} = \Gamma_{i-1} X_{i+1} \quad \text{——} \quad ①$$

Recall

$$X_f := X_i = [x_i \; x_{i+1} \; \cdots \; x_{i+j-1}]$$

$$\Rightarrow \quad X_{i+1} = [x_{i+1} \; x_{i+2} \; \cdots \; x_{i+j}]$$

$$\Gamma_i = - \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-2} \\ \hline CA^{i-1} \end{bmatrix} \begin{array}{l} \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \Gamma_{i-1} \\ \left. \right\} l \text{ rows} \end{array}$$

So $\Gamma^o_{i-1} = \left[U_1 S_1^{1/2} T\right]_{\text{top } l(i-1) \text{ rows}}$

Then, from ①, $X^o_{i+1} = \underbrace{\Gamma^{o\,+}_{i-1}\ \underline{\underline{\xi}}}_{\text{both from data}}$

$$\text{Hence} \quad \overbrace{\begin{bmatrix} X^o_{i+1} \\ Y^o_{i|i} \end{bmatrix}}^{\text{from data}} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X_i^o \\ U_{i|i} \end{bmatrix} \xrightarrow{} \text{from data} \longrightarrow ②$$

$$\left\{ \begin{array}{l|l|l} x_{i+1} = A x_i^o + B u_i^o & x_{i+2}^o = A x_{i+1} + B u_{i+1} & \cdots \\ y_i^o = C x_i^o + D u_i^o & y_{i+1} = C x_{i+1} + D u_{i+1} & \end{array} \right.$$

$$\cdots \quad \left| \begin{array}{l} x_{i+j}^o = A x_{i+j-1} + B u_{i+j-1} \\ y_{i+j-1} = A x_{i+j-1} + B u_{i+j-1} \end{array} \right\}$$

Recall $Y_{i|i} = [y_i \ \ y_{i+1} \ \cdots \ y_{i+j-1}]$

$U_{i|i} = [u_i \ \ u_{i+1} \ \cdots \ u_{i+j-1}]$

Use ② to solve for A, B, C, D.

⇒ for noisy data A, B, C, D has to be solved in least sq. sense

→ without noise, ② is consistent.

<u>Algo 2</u>: Using Γ_i^o -

Define: $\longrightarrow \overline{\Gamma}_i^o$

$$\overline{\Gamma}_i^o \left\{ \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-2} \\ CA^{i-1} \end{bmatrix} \right\} \underline{\Gamma}_i^o$$

Then $\underline{\Gamma}_i^o \cdot A = \overline{\Gamma}_i^o \longrightarrow ③$

② can be solved using LS:

$$A = \underline{\Gamma_i^o}^{+}\,\Gamma_i^o$$

Calculation of C : $C = \left[\Gamma_i^o\right]_{\text{first } l \text{ rows}}$

Calculation of B & D:

1) Compute a full row rank matrix $\Gamma_i^{o\perp} \in \mathbb{R}^{(li^o-n)\times li^o}$ s.t. $\Gamma_i^{o\perp}\Gamma_i^o = 0$

2) Recall: $Y_f = \Gamma_i^o X_f + H_i^o U_f$. Then

$$\Gamma_i^{o\perp} Y_f = \Gamma_i^{o\perp} H_i^o U_f$$

$$\Rightarrow \underbrace{\underbrace{\Gamma_i^{o\perp} Y_f U_f^{\dagger}}_{(li^o-n)\times mi^o}}_{M} = \underbrace{\underbrace{\Gamma_i^{o\perp}}_{(li-n)\times li}}_{L} \underbrace{H_i^o}_{\to\, li^o\times mi^o}$$

$$\underbrace{\begin{bmatrix} \overset{(li^o-n)\times m}{M_1} & M_2 & \cdots & M_{i^o} \end{bmatrix}}_{\text{known}} = \underbrace{\begin{bmatrix} \overset{(li-n)\times l}{L_1} & \cdots & L_{i^o} \end{bmatrix}}_{\text{known}} \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & \ddots & & \vdots \\ \vdots & & \ddots & \\ CA^{i-2}B & \cdots & & D \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_{i^o} \end{bmatrix} = \begin{bmatrix} L_1 & L_2 & \cdots & L_{i-1} & L_i \\ L_2 & L_3 & \cdots & L_i & 0 \\ \vdots & & & & \vdots \\ L_{i^o} & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} I_l & 0 \\ 0 & \underline{\Gamma_1^o} \end{bmatrix} \begin{bmatrix} D \\ B \end{bmatrix}$$

Solve D, B using L.S.

Loose Ends: $[W_p/U_f^{\perp}][W_p/U_f^{\perp}]^{\dagger} W_p = W_p$?

#Two extra assumptions are required:

1> The input is persistently exciting of order $2i$

:= $U_{0/2i-1}$ is full row rank $= 2mi$

[:= $E(U_{0/2i-1}, U_{0/2i-1}^{T})$ is full rank $= 2mi$
— for stochastic case]

2> row sp. of U_f $\cap$ row sp. of X_p
= empty

Claim: rank W_p = rank $W_p/U_f^{\perp}$

Proof: Recall $Y_p = \Gamma_i X_p + H_i U_p$

$$\Rightarrow W_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix} = \begin{bmatrix} I_{mi} & 0 \\ H_i & \Gamma_i \end{bmatrix} \begin{bmatrix} U_p \\ X_p \end{bmatrix}$$

$$\Rightarrow W_p/U_f^{\perp} = \begin{bmatrix} I_{mi} & 0 \\ H_i & \Gamma_i \end{bmatrix} \begin{bmatrix} U_p/U_f^{\perp} \\ X_p/U_f^{\perp} \end{bmatrix}$$

By the assumptions above:

$$\text{rank} \begin{bmatrix} U_p \\ X_p \end{bmatrix} = \text{rank} \begin{bmatrix} U_p/U_f^{\perp} \\ X_p/U_f^{\perp} \end{bmatrix}$$

— Claim proved —

Next denote SVD of $W_p/U_f^{\perp}$.

$$W_p/U_f^{\perp} = [U_1 \;\; U_2]\begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

$$= U_1 S_1 V_1^T$$

By claim above, rank W_p = rank $W_p/U_f^{\perp}$.

Also $W_p/U_f^{\perp} \subset$ col sp of W_p.

$\Rightarrow$ col sp $(W_p/U_f^{\perp})$ = col sp (W_p)

$\Rightarrow$ $W_p = U_1 R$

$R \hookrightarrow$ unknown matrix

Lastly, consider:

$$[W_p/U_f^{\perp}][W_p/U_f^{\perp}]^{\dagger} W_p = [U_1 S_1 V_1^T][V_1 S_1^{-1} U_1^T] U_1 R$$

$$= U_1 \underbrace{S_1 V_1^T V_1 S_1^{-1}}_{I} \underbrace{U_1^T U_1}_{I} R$$

$$= U_1 R = W_p.$$