6.435 Solution set #3 — Abe Elfadel

Problem 1 : ARX model, noise free case

$A(q)\,y = B(q)\,u$; $d^\circ A = n_a$, $d^\circ B = n_b$.

$A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}$; $B(q) = b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}$

$$y(t) = -a_1 y(t-1) - \dots - a_{n_a} y(t-n_a) + b_1 u(t-1) + \dots + b_{n_b} u(t-n_b)$$

$$= \varphi^T(t)\,\theta \quad \text{with}$$

$$\varphi^T(t) = \big(-y(t-1) \;\dots\; -y(t-n_a) \quad u(t-1) \dots u(t-n_b)\big) \in \mathbb{R}^{n_a+n_b}$$

$$\theta = (a_1, \dots a_{n_a}, b_1, \dots, b_{n_b})^T$$

We have $R = E[\varphi(t)\,\varphi^T(t)]$. Note that

$$\varphi^T(t) = \left[-\frac{B(q)}{A(q)} u(t-1), \dots, -\frac{B(q)}{A(q)} u(t-n_a), u(t-1), \dots, u(t-n_b)\right]$$

$$= \frac{1}{A(q)}\big[-B(q)u(t-1), \dots, -B(q)u(t-n_a), A(q)u(t-1), \dots, A(q)u(t-$$

$$= \frac{1}{A(q)}\Big[\big(-b_1 u(t-2) - \dots - b_{n_b} u(t-1-n_b)\big), \dots, \big(-b_1 u(t-1-n_a) \dots - b_{n_b} u(t-$$

$$\big(u(t-1) + a_1 u(t-2) + \dots + a_{n_a} u(t-1-n_a)\big), \dots,$$

$$\big(u(t-n_b) + a_1 u(t-n_b-1) + \dots + a_{n_a} u(t-n_a-n_b)\big)\Big]^T$$

In matrix form we can write :

$$\varphi(t) = \frac{1}{A(q)} \begin{array}{c} \\ \left[\begin{array}{ccccccc} 0 & -b_1 & \dots & -b_{n_b} & & 0 & \\ & 0 & -b_1 & \dots & \dots & -b_{n_b} & \\ 0 & & \ddots & \ddots & & & \\ & & & 0 & -b_1 & \dots & -b_{n_b} \\ 1 & a_1 & \dots & a_{n_a} & & & \\ 0 & 1 & a_1 & \dots & a_{n_a} & 0 & \\ & 0 & & \ddots & & & \end{array}\right] \end{array} \begin{bmatrix} u(t-1) \\ u(t-2) \\ \vdots \\ u(t-n_a) \\ \vdots \\ u(t-n \end{bmatrix} \begin{array}{l} \left.\right\} n_a \\ \\ \left.\right\} n_b \end{array}$$

Therefore $\varphi(t) = \mathcal{S}(-B, A)\,\tilde{\varphi}(t)$ where $\mathcal{S}(-B,A)$ is the Sylvester matrix of the polynomials $-B$ and A, and $\tilde{\varphi}(t) = \frac{1}{A(q)}\left(u(t-1), \ldots, u(t-n_a-n_b)\right)^T \in \mathbb{R}^{n_a+n_b}$

It follows that

$$R = E\left[\varphi(t)\varphi^T(t)\right] = \mathcal{S}(-B,A)\, E\,\tilde{\varphi}(t)\tilde{\varphi}^T(t)\, \mathcal{S}(-B,A)^T$$
$$\triangleq \mathcal{S}(-B,A)\,\tilde{R}\,\mathcal{S}(-B,A)^T.$$

Therefore R is positive definite iff the following two conditions are satisfied:

i) $\mathcal{S}(-B,A)$ is non singular
ii) $\tilde{R}$ is positive definite.

According to the notes distributed in class, $\mathcal{S}(-B,A)$ is nonsingular iff the polynomials A and B are coprime.

Whereas the matrix $\tilde{R} = \frac{1}{[A(q)]^2} E\left\{[u(t-1) \ldots u(t-n_a-n_b)]\,[u(t-1), \ldots, u(t-n_a-n_b)]^T\right\}$

is positive definite is positive definite iff the signal $u(t)$ is persistently exciting of order n_a+n_b. We have therefore the following theorem

Theorem: For the noise free ARX model, the matrix R is positive definite iff A and B are coprime and u is p.e. of order n_a+n_b.

Problem 2 ARX model with noise.

$A(q)\,y = B(q)\,u + e(t)$

A & B as in problem 1, e is WN uncorrelated with u.

A careful look at the proof of problem 1 will reveal that its basis was the expression of $y(t)$ in terms of $u(t)$ in the regression vector $\varphi(t)$.

Here we will proceed similarly by expressing $y(t)$ in terms of both $u(t)$ and the noise process $e(t)$.

We have
$$y(t) = \frac{B(q)}{A(q)} u(t) + \frac{1}{A(q)} e(t) \triangleq \tilde{u}(t) + \tilde{e}(t)$$

Therefore
$$\varphi^T(t) = \left(-y(t-1), \dots, -y(t-n_a), u(t-1), \dots, u(t-n_b)\right)$$

or
$$\varphi^T(t) = \left(-\tilde{u}(t-1) - \tilde{e}(t-1), \dots, -\tilde{u}(t-n_a) - \tilde{e}(t-n_a), u(t-1), \dots, u(t-n_b)\right)$$

which can be written as

$$\varphi^T(t) = \varphi_u^T(t) + \varphi_e^T(t) \quad \text{with}$$

$$\varphi_u^T(t) = \left(-\tilde{u}(t-1), \dots, \tilde{u}(t-n_a), u(t-1), \dots, u(t-n_b)\right) \in \mathbb{R}^{n_a+n_b}$$
$$\varphi_e^T(t) = \left(-\tilde{e}(t-1), \dots, \tilde{e}(t-n_a), 0, \dots, 0\right) \in \mathbb{R}^{n_a}$$

Now
$$R = E\left[\varphi(t)\varphi^T(t)\right] = E\left[(\varphi_u(t) + \varphi_e(t))(\varphi_u^T(t) + \varphi_e^T(t))\right]$$
$$= E\left[\varphi_u(t)\varphi_u^T(t)\right] + E\left[\varphi_e(t)\varphi_e^T(t)\right]$$ because u & are uncorrelated.

So
$$R = E\left[\varphi(t)\varphi^T(t)\right] = \begin{bmatrix} R_{\tilde{u}} & R_{u\tilde{u}} \\ R_{u\tilde{u}}^T & R_u \end{bmatrix} + \begin{bmatrix} R_{\tilde{e}} & 0 \\ 0 & 0 \end{bmatrix}$$

Since the noise $e(t)$ is white the matrix $R_{\tilde{e}}$ is positive definite. With this assumption in mind we want to prove the following

Theorem: R is positive definite iff u is persistently exciting of order n_b (i.e., iff R_u is positive definite).

Proof: ($\Rightarrow$) Assume that $R > 0$. Consider $x^T = (x_1^T, x_2^T)$

where $\underset{\sim}{x}_1 = \underset{\sim}{0} \in \mathbb{R}^{n_a}$ and $\underset{\sim}{0} \neq \underset{\sim}{x}_2 \in \mathbb{R}^{n_b}$. Then

$\underset{\sim}{x}^T R \underset{\sim}{x} = \underset{\sim}{x}_2^T R_u \underset{\sim}{x}_2 > 0$. In other words,

R_u is positive definite, or u is p.e. of order n_b.

($\Leftarrow$) Conversely, assume u p.e. of order n_b. Then R_u is p.d. From lemma A3 (i) (class handout), we conclude that

$$R_{\tilde{u}} - R_{u\tilde{u}} R_u^{-1} R_{u\tilde{u}}^T \geq 0$$

Since the noise is white $R_{\tilde{e}} > 0$ and the rank of R is given by lemma 3 (ii)

$$\text{rank } R = n_b + \text{rank}\left(\underbrace{R_{\tilde{e}}}_{>0} + \underbrace{R_{\tilde{u}} - R_{u\tilde{u}} R_u^{-1} R_{u\tilde{u}}^T}_{\geq 0} \right)$$
$$= n_b + n_a$$

R is therefore a positive semidefinite matrix with full rank. It is therefore p.d. Q.E.D.

Remarks: ① Note the crucial role played by the white noise. It guarantees that $R_{\tilde{e}}$ is of rank n_a

② It would have been sufficient to assume that the filtered noise $\tilde{e}$ is p.e. of order n_a.

Problem 3 Counterexample.

We know that in the scalar case, $u(t)$ is p.e. of order n iff its power spectrum non zero at n distinct frequencies. We know also that in the multivariable case, it is sufficient for the power spectrum to be p.d at n different frequencies for the input to be p.e. of order n. We want to give a counterexample showing that in the multivariable case this condition is not necessary [illegible]

The simplest example I can think of is to take

$\underset{\sim}{u}(t) = (\cos\omega_1 t, \cos\omega_2 t)^T \in \mathbb{C}^2$ with $\omega_1, \omega_2 \in (0, \pi)$ and $\frac{\omega_1}{\omega_2} \notin \mathbb{Q}$.

First this signal is p.e of order at least 2 because the correlation matrix

$$R_u(\tau) = \frac{1}{2}\begin{bmatrix} \cos\omega_1\tau & 0 \\ 0 & \cos\omega_2\tau \end{bmatrix}$$

(see example 2.3 in textbook).

is non singular. However

$$\Phi_u(\omega) = \frac{\pi}{2}\begin{bmatrix} \delta(\omega-\omega_1)+\delta(\omega+\omega_1) & 0 \\ 0 & \delta(\omega-\omega_2)+\delta(\omega+\omega_2) \end{bmatrix}$$

is singular for all frequencies as it might be easily checked.

Problem # 4 Linear regression.

$y(t) = a + bt + e(t)$; $e(t)$ WN with variance λ^2.

1) Estimation of $\theta = (a, b)^T$:

Data are $y(1), \dots, y(N)$.

In this case, $\varphi^T = \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ 1 & 2 & \cdots & N \end{bmatrix}$ and the LS estimate

are given by $(\varphi^T\varphi)\hat{\theta} = \varphi^T Y$ with $Y = (y(1), \dots, y(N))$

To simplify notations, define $S_0 = \sum_{t=1}^{N} y(t)$ & $S_1 = \sum_{t=1}^{N} t\, y(t)$

Then

$$\varphi^T\varphi = \begin{bmatrix} N & \sum_{t=1}^{N} t \\ \sum_{t=1}^{N} t & \sum_{t=1}^{N} t^2 \end{bmatrix} = \begin{bmatrix} N & \frac{N(N+1)}{2} \\ \frac{N(N+1)}{2} & \frac{N(N+1)(2N+1)}{6} \end{bmatrix}$$

$$(\varphi^T\varphi)^{-1} = \begin{bmatrix} \frac{2(2N+1)}{N(N-1)} & \frac{-6}{N(N-1)} \\ \frac{-6}{N(N-1)} & \frac{12}{N(N^2-1)} \end{bmatrix}$$

It follows that :

$$\boxed{\begin{aligned}\hat{a} &= \frac{1}{N(N-1)}\left[2(2N+1)S_0 - 6S_1\right]\\ \hat{b} &= \frac{1}{N(N^2-1)}\left[-6(N+1)S_0 + 12S_1\right]\end{aligned}}$$

(2) Data are $y(-N), y(-N+1), \ldots, y(N)$

Here we denote $S_0 = \sum_{t=-N}^{N} y(t)$, $S_1 = \sum_{t=-N}^{N} t\,y(t)$.

In this case, $\varphi^T = \begin{bmatrix} 1 & 1 & \cdots & & 1 \\ -N & -N+1 & \cdots & 0 & \cdots & N \end{bmatrix}$

So that $\varphi^T\varphi = \begin{bmatrix} 2N+1 & 0 \\ 0 & 2\sum_{t=1}^{N} t^2 \end{bmatrix} = \begin{bmatrix} 2N+1 & 0 \\ 0 & \frac{N(N+1)(2N+1)}{3} \end{bmatrix}$

$$(\varphi^T\varphi)^{-1} = \begin{bmatrix} \frac{1}{2N+1} & 0 \\ 0 & \frac{3}{N(N+1)(2N+1)} \end{bmatrix}$$

So,

$$\boxed{\begin{aligned}\hat{a} &= \frac{S_0}{2N+1}\\ \hat{b} &= \frac{3S_1}{N(N+1)(2N+1)}\end{aligned}}$$

3) Variance of $s(t) = \hat{a} + \hat{b}\,t$

By definition $\operatorname{Var}[s(t)] = E[s(t) - \bar{s}(t)]^2$ where $\bar{s}(t) = E\,s($
But $E\,s(t) = E(\hat{a}) + t\,E(\hat{b})$
$= a + b\,t$, the LS estimate being unbiased
It follows that

$$\operatorname{Var}[s(t)] = E\left[(\hat{a}-a) + (\hat{b}-b)\,t\right]^2$$

$$= E\left[(\hat{a}-a)^2\right] + 2t\,E\left[(\hat{a}-a)(\hat{b}-b)\right] + E\left[(\hat{b}-b\right.$$

i.e., $\mathrm{Var}[s(t)] = \mathrm{Var}(\hat{a}) + 2t\,\mathrm{Cov}(\hat{a},\hat{b}) + t^2\,\mathrm{Var}(\hat{b})$

where $\mathrm{Var}(\hat{a})$, $\mathrm{Var}(\hat{b})$ and $\mathrm{Cov}(\hat{a},\hat{b})$ are given by the coefficients o the covariance matrix of the LS estimates $\hat{a}$ & $\hat{b}$. This matrix is given by

$$\mathrm{Cov}(\hat{\theta}_{LS}) = \lambda^2(\varphi^T\varphi)^{-1}.$$ (See lecture notes on linear regressi

For case 1, we have

$$\mathrm{Var}(\hat{a}) = \frac{2(2N+1)}{N(N-1)}\lambda^2, \quad \mathrm{Var}(\hat{b}) = \frac{12}{N(N^2-1)}\lambda^2, \quad \mathrm{Cov}(\hat{a},\hat{b}) = \frac{-6}{N(N-1)}\lambda$$

Therefore:

$$\boxed{\mathrm{Var}[s(t)] = \frac{2\lambda^2}{N(N-1)}\left[(2N+1) - 6t + \frac{6}{N+1}t^2\right]}$$

Remarks: ① It is interesting to note that

$$\mathrm{Var}[s(1)] = \mathrm{Var}[s(N)] = 2\lambda^2\frac{(2N-1)}{N(N+1)}$$, i.e., the variance of $s(t)$ is the same at the start and the end of the experiment given that all data have been collected.

② There is a time instant at which $\mathrm{Var}[s(t)]$ is minimum. It can be easily checked that this time instant is $\boxed{t_{min} = \frac{N+1}{2}}$. Then we have

$$\boxed{\mathrm{Var}[s(t_{min})] = \frac{\lambda^2}{N}}$$

4) Define the correlation coefficient ρ_N between $\hat{a}$ & $\hat{b}$ by

$$\rho_N = \frac{\mathrm{Cov}(\hat{a},\hat{b})}{\sqrt{\mathrm{Var}(\hat{a})\mathrm{Var}(\hat{b})}} = \frac{-6}{N(N-1)}\sqrt{\frac{N^2(N-1)^2(N+1)}{24(2N+1)}} = -\sqrt{\frac{36(N}{24(2}}$$

So, $$\boxed{\lim \rho_N = -\frac{\sqrt{3}}{2}}$$

Problem 5 : Ljung's exercices :

(a) 4G2 : Colored measurement noise

You have to have the eyes of an eagle to see that there is a calligraphic difference between the input and the output of (4.147).

Model representation
$$x(t+1) = A_1(\theta)x(t) + B_1(\theta)u(t) + w_1(t)$$
$$y(t) = C_1(\theta)x(t) + v(t)$$

Noise representation
$$\xi(t+1) = A_2(\theta)\xi(t) + K(\theta)\mu(t)$$
$$v(t) = C_2(\theta)\xi(t) + \mu(t)$$

$w_1(t)$ is white noise of variance $\overline{R_1}(\theta)$
$\mu(t)$ " " " " " $R_2(\theta)$.

Now we combine the model and noise representations into a single representation by augmenting the state-space. Define

$$z(t) = \begin{bmatrix} x(t) \\ \xi(t) \end{bmatrix} \quad ; \quad A(\theta) = \begin{bmatrix} A_1(\theta) & 0 \\ 0 & A_2(\theta) \end{bmatrix} \quad ; \quad B(\theta) = \begin{bmatrix} B_1(\theta) \\ 0 \end{bmatrix}$$

$$C(\theta) = \begin{bmatrix} C_1(\theta) & C_2(\theta) \end{bmatrix} \quad ; \quad w(t) = \begin{bmatrix} w_1(t) \\ K(\theta)\mu(t) \end{bmatrix}$$

Then we have the state-space representation:

$$\begin{cases} z(t+1) = A(\theta)\, z(t) + B(\theta)\, u(t) + w(t) \\ y(t) = C(\theta)\, z(t) + \mu(t) \end{cases}$$

The covariance matrices of the noise processes in this model are given by

$$\begin{cases} R_1(\theta) = E\left[w(t)\, w^T(t)\right] = \begin{bmatrix} \overline{R_1(\theta)} & 0 \\ 0 & K(\theta)R_2(\theta)K^T(\theta) \end{bmatrix} \\ R_{12}(\theta) = E\left[w(t)\, \mu^T(t)\right] = \begin{bmatrix} 0 \\ K(\theta)R_2(\theta) \end{bmatrix} \\ R_2(\theta) = E\left[\mu(t)\, \mu^T(t)\right] \end{cases}$$

(b) 4.6.3

Verification of the steady state Kalman Filter:

* Output noise of the state-space model:

$v_1(t) = C(qI-A)^{-1}w(t) + v(t)$ with $E\,w(t)\,w^T(t) = R_1$

and $E[v^2(t)] = R_2$. Note that R_2 is a scalar.

The spectrum of $v_1(t)$ is

$$\Phi_1(\omega) = C(e^{i\omega}I-A)^{-1}R_1(e^{-i\omega}I-A)^{-T}C^T + R_2$$

* Output noise of the innovation model is

$v_2(t) = (C(qI-A)^{-1}K + I)\,e(t)$ with $E[e^2(t)] = \lambda$

The spectrum of $v_2(t)$ is

$$\Phi_2(\omega) = \lambda\left[C(e^{i\omega}I-A)^{-1}K+1\right]\left[C(e^{-i\omega}I-A)^{-1}K+1\right]^T.$$

The equations to remember here are:

Innovation variance $\lambda = CPC^T + R_2$ (This is a scalar) (4.88b)

Kalman gain $K = APC^T(CPC^T+R_2)^{-1} = \frac{1}{\lambda}APC^T$ (4.84a)

Ricatti equation (steady state) $P = APA^T + R_1 - \lambda KK^T$ (4.84b)

(a) To prove that $\Phi_1(\omega) \equiv \Phi_2(\omega)$, let us form the difference

$$\begin{aligned}\Phi_1(\omega)-\Phi_2(\omega) &= C(e^{i\omega}I-A)^{-1}R_1(e^{-i\omega}I-A)^{-1}C^T + R_2 \\ &\quad - \lambda\big[C(e^{i\omega}I-A)^{-1}KK^T(e^{-i\omega}I-A)^{-1}C^T + 1 \\ &\quad + C(e^{i\omega}I-A)^{-1}K + K^T(e^{-i\omega}I-A)^{-1}C^T\big] \\ &= C(e^{i\omega}I-A)^{-1}(R_1-\lambda KK^T)(e^{-i\omega}I-A)^{-1}C^T \\ &\quad + \underbrace{R_2-\lambda}_{-CPC^T} - C(e^{i\omega}I-A)^{-1}\underbrace{\lambda K}_{APC^T} - \underbrace{\lambda K^T}_{CPA^T}(e^{-i\omega}I-A)^{-1} \\ &= C(e^{i\omega}I-A)^{-1}\big[P-APA^T-(e^{i\omega}I-A)P(e^{-i\omega}I-A) \\ &\quad -AP(e^{-i\omega}I-A)-(e^{i\omega}I-A)PA^T\big](e^{-i\omega}I-A)^{-1}C^T\end{aligned}$$

Now, it is easy to verify that what is inside the square brackets is 0

Remark: We have proved, in the special case of $\dim y = 1$, that the steady-state Kalman filter defines a spectral factorization for the state space model (put $B=0$). The spectral factor is given by

$$W(z) = C(zI-A)^{-1}K + 1.$$

This result is valid for a multivariable system. A general proof can be found in Anderson & Moore (Section 4.5).

(b) We have $H(q) = C(qI-A)^{-1}K + 1$. We are asked to prove that

$$1 - H^{-1}(q) = C(qI-A+KC)^{-1}K.$$

I am going to start from the right hand side because of something called the matrix inversion lemma:

$$C(qI-A+KC)^{-1}K = C\left\{(qI-A)^{-1} - (qI-A)^{-1}K\left[1+C(qI-A)^{-1}K\right]^{-1} C(qI-A)^{-1}\right\}K$$

$$= C(qI-A)^{-1}K - C(qI-A)^{-1}K\left[1+C(qI-A)^{-1}K\right]^{-1} C(qI-A)^{-1}K$$

$$= C(qI-A)^{-1}K - \frac{C(qI-A)^{-1}KC(qI-A)^{-1}K}{1+C(qI-A)^{-1}K}$$

$$= \frac{C(qI-A)^{-1}K}{1+C(qI-A)^{-1}K} = 1 - \left[1+C(qI-A)^{-1}K\right]^{-1}$$

i.e.

$$\boxed{C(qI-A+KC)^{-1} = 1 - H^{-1}(q)}$$

As for the second formula, we have

$$C(qI-A+KC)^{-1}B = C(qI-A)^{-1}B - \frac{C(qI-A)^{-1}KC(qI-A)^{-1}B}{1+C(qI-A)^{-1}K}$$

$$= \ldots 1 + C(qI-A)^{-1}K \quad C(qI-A)^{-1}B$$

Remark: ... have proved ...

c) Using the formulas derived above, we get

$$\hat{y}(t|\theta) = H^{-1}(q,\theta) G(q,\theta) u(t) + [1 - H^{-1}(q,\theta)] y(t) \quad (3.20)$$

$$= C(qI - A + KC)^{-1} B u(t) + C(qI - A + KC)^{-1} K y(t)$$

(see 4.85)

Which is the steady state Kalman filter for this special case.

(c) 4G4 : local identifiability

let $W(z,\theta)$ be a model structure with gradient $\psi(z,\theta) = \frac{d}{d\theta} W(z,\theta)$

Define $\Gamma_1(\theta) = \int_{-\pi}^{\pi} \psi(e^{i\omega},\theta)\, \psi^T(e^{-i\omega},\theta)\, d\omega$

(a) If $\Gamma_1(\theta)$ is nonsingular then $\mathcal{M}$ is locally identifiable.

let θ be a parameter and let θ_1 be close enough to θ_1.
Then using a Taylor series expansion we get

$$W(z,\theta_1) = W(z,\theta) + (\theta_1 - \theta)\psi(z,\theta) + o(|\theta_1 - \theta|)$$

small o

where $\lim_{\theta_1 \to \theta} \frac{o(\|\theta_1 - \theta\|)}{\|\theta_1 - \theta\|} = 0$

Now assume $W(z,\theta_1) = W(z,\theta)$. Then

$$0 = (\theta_1 - \theta)\psi(z,\theta) + o(\|\theta_1 - \theta\|) \quad . \text{ Similarly}$$

$$0 = \psi^T(z,\theta)(\theta_1 - \theta)^T + o(\|\theta_1 - \theta\|).$$

Therefore $(\theta_1 - \theta)\, \psi(e^{i\omega},\theta)\, \psi^T(e^{i\omega},\theta)(\theta_1 - \theta)^T = o(\|\theta_1 - \theta\|^2)$

Integrating with respect to ω we get

$$(\theta_1 - \theta)\, \Gamma_1(\theta)\, (\theta_1 - \theta)^T = 2\pi\, o(\|\theta_1 - \theta\|^2)$$

definite. This means that the above equality is valid only if $\theta = \theta_1$. According to definition 4.128 in textbook, this proves local identifiability at θ.

(b) Define $T'(e^{i\omega}, \theta) = \frac{d}{d\theta} T(e^{i\omega}, \theta)$

$$= \frac{d}{d\theta}\left[\frac{d}{d\theta} G(e^{i\omega}, \theta) \quad \frac{d}{d\theta} H(e^{i\omega}, \theta)\right] \quad (4.122$$

But (see 4.121)

$$\psi(e^{i\omega}, \theta) = \frac{1}{(H(e^{i\omega}, \theta))^2} T'(e^{i\omega}, \theta) \begin{bmatrix} H(e^{i\omega}, \theta) & 0 \\ -G(e^{i\omega}, \theta) & 1 \end{bmatrix}$$

Now the matrix $\frac{1}{(H(e^{i\omega}, \theta))^2} \begin{bmatrix} H(e^{i\omega}, \theta) & 0 \\ -G(e^{i\omega}, \theta) & 1 \end{bmatrix}$ has a determinant equal $\frac{1}{H(e^{i\omega}, \theta)}$ and therefore is nonsingular. It follows that $\psi(e^{i\omega}, \theta)$ and $T'(e^{i\omega}, \theta)$ have the same rank $\forall \omega \in [-\pi, \pi]$.

Consequently $\Gamma_2(\theta) = \int_{-\pi}^{\pi} T'(e^{i\omega}, \theta)\left[T'(e^{-i\omega}, \theta)\right]^T d\omega$ is nonsingular if and only if $\Gamma_1(\theta)$ is.

(d) 4G.8 : δ- parametrization.

I am going to treat the general case. Consider the δ-ARX model

$\bar{A}(\delta) y(t) = \bar{B}(\delta) u(t) + e(t)$ where $\delta = 1 - q^{-1}$. Now

$$\bar{A}(\delta) = \bar{a}_{n_a} + \cdots + \bar{a}_1 \delta^{n_a - 1} + \delta^{n_a} = \sum_{k=0}^{n_a} \bar{a}_k \delta^{n_a - k}$$

$$= \sum_{k=0}^{n_a} \bar{a}_k \sum_{j=0}^{n_a - k} \binom{n_a - k}{j} (-1)^j q^{-j}$$

$$= \sum_{j=0}^{n_a} (-1)^j \left\{ \sum_{k=0}^{n_a - j} \bar{a}_k \binom{n_a - k}{j} \right\} q^{-j} =$$

In terms of q^{-1}, we have $A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}$.
Identifying $\bar{A}(\delta)$ and $A(q)$ term by term gives:

$$\boxed{a_k = (-1)^k \sum_{\ell=0}^{n_a-k} \bar{a}_\ell \binom{n_a-\ell}{k}}$$

Which is the sought reparametrization of $A(q)$. Clearly, we get a similar formula for $B(q)$.

It should be intuitively clear why the δ parametrization is better for short sampling interval T. Because of sampling the poles of the analog system are mapped by $\exp(sT)$ to the z-plane. When T is short, these poles, including the stable ones, will cluster around the point $(1,0)$. Therefore any numerical error might make the sampled system unstable. In the $\delta = 1 - q^{-1}$ parametrization everything is shifted back to the center of the unit circle.

(c) 4E6 Identifiability:

Transfer functions of the θ-model

$$G(q,\theta) = \frac{b_1 q + b_2}{q^2 + a_1 q + a_2} \qquad H(q,\theta) = \frac{k_1 q + k_2}{q^2 + a_2 q + a_1} + 1$$

Transfer functions of the η-model

$$G(q,\eta) = \frac{(\gamma_1\mu_1 + \gamma_2\mu_2) q - (\gamma_1\mu_1\lambda_2 + \gamma_2\mu_2\lambda_1)}{q^2 - (\lambda_1+\lambda_2) q + \lambda_1\lambda_2}$$

$$H(q,\eta) = \frac{(\gamma_1\bar{k}_1 + \gamma_2\bar{k}_2) q - (\gamma_1\bar{k}_1\lambda_2 + \gamma_2\bar{k}_2\lambda_1)}{q^2 - (\lambda_1+\lambda_2) q + \lambda_1\lambda_2} + 1$$

Obviously: $D_\eta = \mathbb{R}^8$.

For the two models do describe the same model set, the θ & η transfer functions must be made identical. Identifying the polynomial coefficients we get

$a_1 = -\lambda_1 - \lambda_2$ (1)

$a_2 = \lambda_1 \lambda_2$ (2)

$b_1 = \gamma_1 \mu_1 + \gamma_2 \mu_2$ (3)

$b_2 = -\gamma_1 \mu_1 \lambda_2 - \gamma_2 \mu_2 \lambda_1$ (4)

$k_1 = \gamma_1 \bar{k}_1 + \gamma_2 \bar{k}_2$ (5)

$k_2 = -\gamma_1 \bar{k}_1 \lambda_2 - \gamma_2 \bar{k}_2 \lambda_1$ (6)

We determine λ_1, λ_2, $\gamma_1 \mu_1$, $\gamma_2 \mu_2$, $\gamma_1 \bar{k}_1$ & $\gamma_2 \bar{k}_2$ from θ as follows:

We use (1) and (2) to get λ_1 & λ_2. Then we solve the two linear systems (3), (4) & (5), (6) to get the remaining parameters. Note that the η-model is very redundant. $\gamma_i \mu_i$, $\gamma_i \bar{k}_i$, $i = 1, 2$, can be replaced with $m\gamma_i \frac{\mu_i}{m}$, $m\gamma_i \frac{\bar{k}_i}{m}$, $m \neq 0$, without changing the model. Therefore the η-model is not identifiable.

As for the θ-model which is in canonical observable form (see 4.143), it is identifiable iff the matrix
$[A(\theta), (B(\theta)\ K(\theta))]$ is controllable.

Remark: The set D_1 for θ is defined by the equations (1) through (6) above. However one can be more specific given that we impose on λ_1 & λ_2 to be real. For this to happen the equation
$\lambda_1^2 + a_1 \lambda_1 + a_2 = 0$ (from (1) & (2)) must always have a real solution. This occurs only if $\boxed{a_1^2 - 4a_2 \geq 0.}$

Therefore D_1 is defined by the range of the mapping. Although intersected with $\{\theta \in \mathbb{R}^6 \mid a_1^2 - 4a_2 \geq 0\}$