LUNDS TEKNISKA
HÖGSKOLA
Lunds universitet

Institutionen för
REGLERTEKNIK

FRT 041 System Identification

Final Exam March 13, 2014, 8am - 1pm

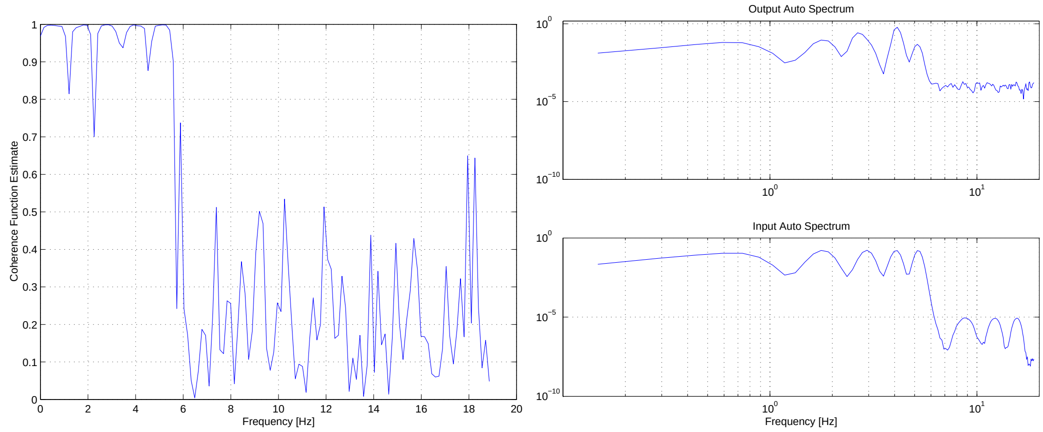
General Instructions

This is an open book exam. You may use any book you want, including the slides from the lecture, but no exercises, exams, or solution manuals are allowed. Solutions and answers to the problems should be well motivated. The exam consists of 8 problems. The credit for each problem is indicated in the problem. The total number of credits is 25 points. Preliminary grade limits:

Grade 3:	12 – 16 points
Grade 4:	17 – 21 points
Grade 5:	22 – 25 points

Results

The results of the exam will be posted at the latest March 21, 2014 on the note board on the first floor of the M-building.



Figur 1 Left: coherence spectrum. Right: input and output autospectra. Notice that the frequency axis is linear in the coherence plot, but logarithmic in the spectrum plots.

1. In Figure 1 the coherence spectrum as well as the input and output auto spectra for a particular experiment are shown. The sampling frequency used when obtaining the data was 37.71 Hz.

a. In what frequency range can a model estimated from this data be expected to describe the true system accurately? Motivate your answer. (1 p)

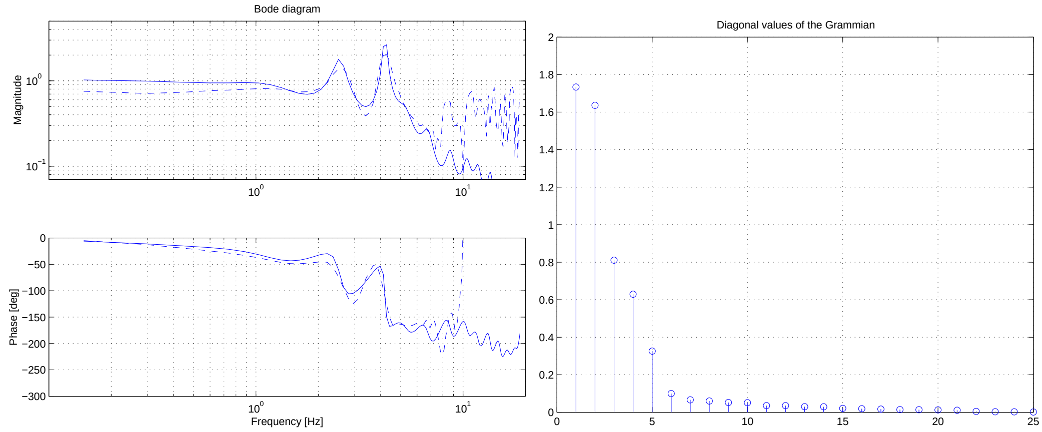
b. What can be said about the choice of sampling frequency for the experiment? Discuss the potential risks of choosing too high respectively too low sampling frequency when doing system identification! (2 p)

c. The choice of model order is an important and often difficult problem in system identification. In this problem, two different models has been estimated from the data used in problem *a*. The first model consists of an estimate of the transfer function, which has been obtained using the Matlab command `spa`. A Bode plot for the model is shown in Figure 2 (left). The second model is a 25th-order ARX model, which has been obtained using the Matlab command `arx`. The Bode plot for this model is also shown in Figure 2. The ARX model was converted to a state-space model, and a balanced realization was calculated. The diagonal values of the Gramian are shown in Figure 2 (right).

Your task is to suggest a suitable model order for the system. You answer should be well motivated. (2 p)

Solution

a. The coherence spectrum is fairly close to one in the frequency range 0 Hz to 5.5 Hz, which indicates that the system that generated the data may be well described by a linear system model in this range. For frequencies above 5.5 Hz the coherence spectrum is close to zero, which indicates that the data are not suitable for estimating linear models describing the system for high frequencies. By studying the input autospectrum it may be concluded that



Figur 2 Left: Bode diagrams for the transfer function estimation (dashed) and ARMAX model (solid) respectively. Right: Diagonal values of the Gramian for the balanced realization.

the energy content in the input signal is very small for frequencies above 5.5 Hz, which could be a partial explanation to this observation.

b. As a rule of thumb, a reasonable way of choosing the sampling interval, h, is to let

$$\omega h = [0.2, 0.6]$$

where ω represents important frequencies of the system, such as the cross over frequency or the natural frequency. In our case, we cannot expect to obtain a model valid for frequencies above 5.5 Hz, which corresponds to $\omega h \approx 0.9$. However, in order to fully judge the choice of sampling interval we must have knowledge about the significant frequencies of the system.

In general, if the sampling interval is chosen very short relative to the significant frequencies of the system, this could lead to numerical precision problems. On the other hand, by choosing a too long sampling interval, there is a risk that important dynamics above the Nyquist frequency are not described by the resulting model.

c. The plot of the diagonal values of the Gramian suggests a 5th-order model since the first five values are significantly larger than the others. Also, from the Bode diagram, it seems like the system has two resonance peaks, which indicates that the system have at least four poles.

2. You are trying to estimate the parameters from the moving average process

$$y(k) = au(k-1) + bu(k-3) + e(k). \tag{1}$$

where $\{e(k)\}$ is a zero mean white noise process with variance σ^2 and $\{u(k)\}$ is a zero mean weakly stationary process with autocovariance function $C_{uu}(\tau) = (1/2)^{|\tau|}$ that is uncorrelated with $\{e(k)\}$.

We are interested in finding the least squares estimate for $\hat{\theta} = (\hat{a} \quad \hat{b})^T$. Does the parameter estimate have an asymptotic distribution? If so, what is the distribution and its parameters? (4 p)

Solution

The model (1) can be written as $y(k) = \varphi^T\theta + e(k)$, where $\theta = \begin{pmatrix} a & b \end{pmatrix}$ and

$$\varphi(k)^T = \begin{pmatrix} u(k-1) & u(k-3) \end{pmatrix}.$$

Given N samples of observed input data the regression matrix is

$$\Phi_N = \begin{pmatrix} u(3) & u(1) \\ u(4) & u(2) \\ \vdots & \vdots \\ u(N+2) & u(N) \end{pmatrix}.$$

Now

$$\frac{1}{N}\,\Phi_N^T\,\Phi_N = \frac{1}{N}\begin{pmatrix} \sum_{k=3}^{N+2} u(k)^2 & \sum_{k=3}^{N+2} u(k)u(k-2) \\ \sum_{k=3}^{N+2} u(k)u(k-2) & \sum_{k=1}^{N} u(k)^2 \end{pmatrix}$$

so, under ergodic conditions,

$$E(\frac{1}{N}\,\Phi_N^T\,\Phi_N) = \lim_{N\to\infty}\frac{1}{N}\,\Phi_N^T\,\Phi_N = \begin{pmatrix} C_{uu}(0) & C_{uu}(2) \\ C_{uu}(2) & C_{uu}(0) \end{pmatrix} = \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}$$

Thus the regression matrix is invertible when the number of samples goes to infinity. This fact together with the fact that the input signal is uncorrelated with the noise signal ensures a consistent estimate. Therefore $E(\hat\theta) = \theta$ and the central limit theorem (6.98) on page 121 in the book gives the asymptotic distribution

$$\hat\theta \sim \mathrm{AsN}(\theta, \frac{\sigma^2}{N}\Sigma)$$

where

$$\Sigma = \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}^{-1} = \frac{16}{15}\begin{pmatrix} 1 & -0.25 \\ -0.25 & 1 \end{pmatrix}$$

3. Devise a recursive algorithm to identify a general ARX process

$$y(k) = G_\theta(z^{-1})u(k) + e(k)$$

where $\{e(k)\}$ is white noise with variance 1 uncorrelated with $\{u(k)\}$. Given that we start from y_1 and u_1, what is the first y_k in the residual calculation $(\epsilon_k = y_k - \varphi_k\hat\theta_{k-1})$? (2 p)

Solution

Postulate a transfer function

$$G_\theta(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}}.$$

This gives the difference equation

$$\begin{aligned} &y(k) + a_1 y(k-1) + a_2 y(k-2) + \cdots + a_{n_a} y(k-n_a) \\ &= b_1 u(k-1) + b_2 u(k-2) + \cdots + b_{n_b} u(k-n_b) + e(k) \end{aligned}$$

and thus the regression model

$$y(k) = \varphi(k)^T \theta + e(k)$$

with

$$\varphi(k) = \begin{pmatrix} -y(k-1) \\ -y(k-2) \\ \vdots \\ -y(k-n_a) \\ u(k-1) \\ u(k-2) \\ \vdots \\ u(k-n_b) \end{pmatrix}$$

and

$$\theta = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n_a} & b_1 & b_2 & \cdots b_{n_b} \end{pmatrix}^T$$

Recursive least squares can now be used directly. The first observation will be $y(\max(n_b, n_a) + 1)$.

4. Consider a linear system

$$y(k) = G(z^{-1})u(k) + v(k) \tag{2}$$

where $\{u(k)\}$ and $\{v(k)\}$ are uncorrelated. Given estimates $S_{uu}(e^{i\omega})$ and $S_{yy}(e^{i\omega})$ of the input and output spectrum respectively find an estimate of the disturbance spectrum $S_{vv}(e^{i\omega})$ such that the estimate is dependent on the quadratic coherence spectrum between y and u. This dependency must be shown. (2 p)

Hint: the quadratic coherence spectrum

$$\gamma_{xy}^2(\omega) = \frac{|S_{xy}(e^{i\omega})|^2}{S_{xx}(e^{i\omega})S_{yy}(e^{i\omega})}$$

Solution

The signal model (2) directly gives

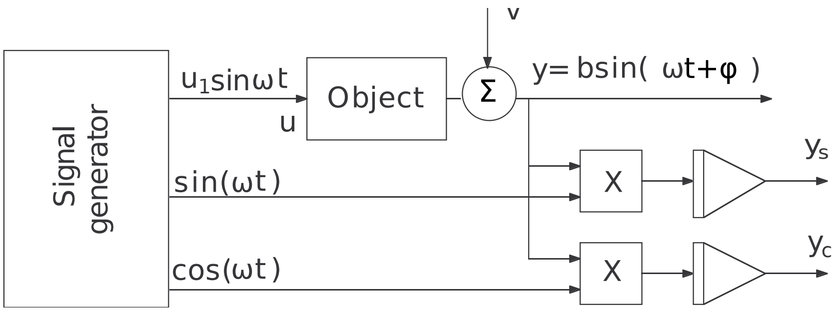
$$\begin{aligned} S_{yu} &= G(e^{i\omega})S_{uu}(e^{i\omega}) \\ S_{yy} &= |G(e^{i\omega})|^2 S_{uu}(e^{i\omega}) + S_{vv}(e^{i\omega}) \end{aligned}$$

The natural estimate is thus

$$\begin{aligned} \hat{S}_{vv} &= \hat{S}_{yy}(e^{i\omega}) - |G(e^{i\omega})|^2 \hat{S}_{uu}(e^{i\omega}) \\ &= \hat{S}_{yy}(e^{i\omega}) - \frac{|G(e^{i\omega})|^2 |\hat{S}_{uu}(e^{i\omega})|^2}{\overline{\hat{S}_{uu}}(e^{i\omega})} \\ &= \hat{S}_{yy}(e^{i\omega}) - \frac{|\hat{S}_{yu}(e^{i\omega})|^2}{\overline{\hat{S}_{uu}}(e^{i\omega})} \\ &= \hat{S}_{yy}(e^{i\omega}) \left(1 - \underbrace{\frac{|\hat{S}_{yu}(e^{i\omega})|^2}{\hat{S}_{uu}(e^{i\omega})\hat{S}_{yy}(e^{i\omega})}}_{\text{Coherence spectrum}} \right) \end{aligned}$$

In the final step, note $\hat{S}_{uu}(e^{i\omega})$ is real since $C_{uu}(\tau)$ is symmetric.

5. Consider the setup for frequency analysis in Figure 3 with $s_T = y_s$ $c_T = y_c$. Assume the disturbance v affects the output of the system. Also assume that the measurement time is T for each test signal with frequency ω_i. How will the estimated frequency response be affected when the disturbance v is



Figur 3 Noise corrupted frequency response analysis.

a. a constant i.e., $v(t) = v_0$? (1 p)

b. 'white' noise with mean 0 and variance σ^2? How much is it necessary to increase the measurement duration T or the gain of the input signal u_1 in order to reduce the variance of the estimated frequency response, $\hat{G}$, by a factor 4? (3 p)

Solution

Let $\phi(\omega) = \arg\{G(i\omega)\}$.

The output from the sine channel is given by,

$$\begin{aligned} s_T =& \int_0^T (u_1|G(i\omega)|\sin(\omega t + \phi(\omega)) + v(t))\sin\omega t dt \\ =& \dots \\ =& \frac{1}{2}u_1 T|G(i\omega)|\cos\phi(\omega) + \triangle s_T \end{aligned}$$

where

$$\triangle s_T = \int_0^T v(t)\sin\omega t dt.$$

The output of the cosine channel is given by,

$$\begin{aligned} c_T =& \int_0^T (u_1|G(i\omega)|\sin(\omega t + \phi(\omega)) + v(t))\cos\omega t dt \\ =& \dots \\ =& \frac{1}{2}u_1 T|G(i\omega)|\sin\phi(\omega) + \triangle c_T \end{aligned}$$

where

$$\triangle c_T = \int_0^T v(t)\cos\omega t dt.$$

The estimate of $G(i\omega)$ becomes,

$$\hat{G}(i\omega) = G(i\omega) + \triangle G(i\omega)$$

where the error $\triangle G$ is

$$\triangle G(i\omega) = \frac{2}{Tu_1}(\triangle s_T + i \triangle c_T)$$

a. In the case of a constant disturbance $v(t) = v_0$, we get

$$\begin{aligned}\triangle s_T &= 0,\\ \triangle c_T &= 0.\end{aligned}$$

The error due to a constant disturbance is thus zero.

b. We have

$$\hat{G}(i\omega) = G(i\omega) + \triangle G(i\omega)$$

The accuracy of $\hat{G}$ is then determined by the statistical properties of $\triangle G$. The mean value of $\triangle G$ is

$$\mathcal{E}\{\triangle G\} = \frac{2}{u_1 T}\left(\mathcal{E}\{\triangle s_T\} + i\mathcal{E}\{\triangle c_T\}\right)$$

where

$$\begin{aligned}\mathcal{E}\{\triangle s_T\} &= \int_0^T \mathcal{E}\{v(t)\} \sin \omega t dt\\ &= 0\end{aligned}$$

and

$$\begin{aligned}\mathcal{E}\{\triangle c_T\} &= \int_0^T \mathcal{E}\{v(t)\} \cos \omega t dt\\ &= 0\end{aligned}$$

Thus $\mathcal{E}\{\triangle G\} = 0$ which implies that $\mathcal{E}\{\hat{G}\} = \mathcal{E}\{G\}$.
The variance properties of $\hat{G}$ is given by

$$\text{Var}\{\hat{G}(i\omega)\} = \text{Var}\{\triangle G(i\omega)\}.$$

The variance of $\triangle G$ is given by

$$\text{Var}\{\triangle G\} = \text{Var}\{\text{Re } \triangle G\} + Var\{\text{Im } \triangle G\}.$$

where

$$\begin{aligned}\text{Var}\{\text{Re } \triangle G\} &= \mathcal{E}\left\{\frac{2}{u_1 T}\int_0^T v(t) \sin \omega t dt \cdot \frac{2}{u_1 T}\int_0^T v(t) \sin \omega s ds\right\}\\ &= \frac{4\sigma^2}{2u_1^2 T}\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}\{\mathrm{Im}\ \triangle\ G\} &= \mathcal{E}\left\{\frac{2}{u_1T}\int_0^T v(t)\cos\omega t dt\cdot\frac{2}{u_1T}\int_0^T v(t)\cos\omega s ds\right\}\\ &= \frac{2\sigma^2}{u_1^2T}\end{aligned}$$

Thus

$$\mathrm{Var}\{\triangle G\} = \frac{4\sigma^2}{2u_1^2T} + \frac{2\sigma^2}{u_1^2T} = \frac{4\sigma^2}{u_1^2T}$$

To reduce the variance with a factor 4 we must either increase the integration time from T to $4T$, or increase the gain of the input signal from u_1 to $2u_1$.

6. Consider the system

$$m\ddot{q} = -kq + \tau \tag{3}$$

Show that using operator $\lambda = 1/(1+s\tau_0)$, it is possible to identify parameter m and k when $\ddot{q}$ is available. (2 p)

Solution

Application of λ to Eq. 3 gives

$$m\frac{d}{dt}(\lambda\{\dot{q}\}) = -k\lambda\{q\} + \lambda\{\tau\} \tag{4}$$

From the operator algebra we have

$$p\lambda = \frac{p}{1+\tau_0 p} = \frac{1}{\tau_0}(1 - \frac{1}{1+\tau_0 p}) = \frac{1}{\tau_0}(1-\lambda)$$

Thus we find

$$\frac{d}{dt}(\lambda\{\dot{q}\}) = \frac{1}{\tau_0}\dot{q} - \frac{1}{\tau_0}\lambda\{\dot{q}\}$$

And Eq. 4 simplifies to

$$\lambda\{\tau\} = \frac{1}{\tau_0}m(\dot{q} - \lambda\{\dot{q}\}) + k\lambda\{q\} = \frac{1}{\tau_0}\lambda\{\dot{q}\} = \phi^T\theta$$

where

$$\begin{aligned}\phi &= (\,(\dot{q} - \lambda\{\dot{q}\})/\tau_0 \quad \lambda\{q\}\,)^T\\ \theta &= (\,m \quad k\,)^T\end{aligned}$$

7. Consider the discrete transfer function

$$H(z) = \frac{0.75z + 0.35}{z^2 + 0.5z}$$

a. Show that

$$\begin{aligned}\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} &= \begin{bmatrix} -0.0348 & 0.1272 \\ 0.1272 & -0.4652 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 0.8660 \\ 0.0049 \end{bmatrix} u(k) \\ y(k) &= \begin{bmatrix} 0.8660 & 0.0049 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}\end{aligned}$$

is a state-space realization of $H(z)$ (actually balanced). Determine the asymptotic reachability Gramian P and the asymptotic observability Gramian Q. (2 p)

b. Given the Gramians in **a.** investigate if a model-order reduction is possible or not. If such a reduction of $H(z)$ is possible and advisable, find the reduced-order model. (2 p)

Solution

a. For the given state-space realization $\{\Phi, \Gamma, C\}$, direct calculations give

$$C(zI - \Phi)^{-1}\Gamma = H(z)$$

For a balanced realization, the asymptotic reachability Gramian P is equal to the asymptotic observability Gramian Q. The diagonal matrix $\Sigma = P = Q$ fulfills the discrete-time Lyapunov equations

$$\begin{aligned} P &= \Phi P \Phi^T + \Gamma\Gamma^T \\ Q &= \Phi^T Q \Phi + C^T C \end{aligned}$$

Solving the first equation gives

$$\begin{aligned}\begin{bmatrix} p_{11} & 0 \\ 0 & p_{22} \end{bmatrix} &= \begin{bmatrix} -0.0348 & 0.1272 \\ 0.1272 & -0.4652 \end{bmatrix} \begin{bmatrix} p_{11} & 0 \\ 0 & p_{22} \end{bmatrix} \begin{bmatrix} -0.0348 & 0.1272 \\ 0.1272 & -0.4652 \end{bmatrix} \\ &+ \begin{bmatrix} 0.8660 \\ 0.0049 \end{bmatrix} \begin{bmatrix} 0.8660 \\ 0.0049 \end{bmatrix}^T\end{aligned}$$

The solution is

$$\Sigma = P = Q = \begin{bmatrix} 0.7511 & 0 \\ 0 & 0.0155 \end{bmatrix} \tag{5}$$

Since $\Phi = \Phi^T$ and $C = \Gamma^T$ the second Lyapunov equation gives the same result.

b. Yes, since the eigenvalues of the Gramians are not of the same order, $\sigma_1 \gg \sigma_2$ it is advisable to reduce the model order.

We can expect the reduced order model to describe the system "well". A truncation of x_2 (see the book p.219) yields

$$\begin{aligned}
x_1(k+1) &= \left(-0.0348 + \frac{0.1272^2}{1+0.4652}\right) x_1(k) + \left(0.8660 + \frac{0.1272 \cdot 0.0049}{1+0.4652}\right) u(k) \\
&= -0.0237 x_1(k) + 0.8664 u(k)
\end{aligned}$$

$$\begin{aligned}
y(k) &= \left(0.8660 + \frac{0.0049 \cdot 0.1272}{1+0.4652}\right) x_1(k) + \frac{0.0049^2}{1+0.4652} u(k) \\
&= 0.8664 x_1(k) + 0.000016 u_k
\end{aligned}$$

The reduced order transfer-function becomes

$$H_2(z) = \frac{0.000016z + 0.7507}{z + 0.02372}$$

Sometimes it is a concern if this transfer function does not have the same static gain as $H(z)$. Therefore we may compensate for this and obtain:

$$H_3(z) = \frac{H_2(z)}{H_2(e^0)} = \frac{0.000022z + 1.024}{z + 0.02372}$$

8. The impulse response coefficients (or Markov parameters) $\{h_k\}_{k=1}^{\infty}$ form the transfer function

$$H(z) = \sum_{k=1}^{\infty} h_k z^{-k}, \quad h_k = CA^{k-1}B$$

a. Show that a Hankel matrix of these coefficients can be factorised as

$$\begin{aligned}
\mathcal{H}_{r,s}^{(k)} &= \begin{pmatrix} h_{k+1} & h_{k+2} & \cdots & h_{k+s} \\ h_{k+2} & h_{k+3} & \cdots & h_{k+s+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+r} & h_{k+r+1} & \cdots & h_{k+r+s-1} \end{pmatrix} \\
&= \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{pmatrix} A^k \begin{pmatrix} B & AB & \ldots & A^{s-1}B \end{pmatrix}
\end{aligned}$$

(1 p)

b. How can this fact be exploited for system identification purposes? (1 p)

Solution

a. One way is to verify the factorization property by direct substitution of Markov parameters $h_k = CA^{k-1}B$ into the Hankel matrix.

$$
\begin{aligned}
\mathcal{H}_{r,s}^{(k)} &= \begin{pmatrix} h_{k+1} & h_{k+2} & \cdots & h_{k+s} \\ h_{k+2} & h_{k+3} & \cdots & h_{k+s+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+r} & h_{k+r+1} & \cdots & h_{k+r+s-1} \end{pmatrix} \\
&= \begin{pmatrix} CA^kB & CA^{k+1}B & \cdots & CA^{k+s-1}B \\ CA^{k+1}B & CA^{k+2}B & \cdots & CA^{k+s}B \\ \vdots & \vdots & \ddots & \vdots \\ CA^{k+r-1}B & CA^{k+r}B & \cdots & CA^{k+r+s+2}B \end{pmatrix} \\
&= \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{pmatrix} A^k \begin{pmatrix} B & AB & \ldots & A^{s-1}B \end{pmatrix}
\end{aligned}
$$

b. Using a numerical factorization such as the singular value decomposition it is possible to find estimates of the extended observability and controllability matrices. In turn, this information can be used to determine a state-space realization $\{A, B, C\}$. In the factorization above, the matrix

$$
\mathcal{O}_r = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{pmatrix}
$$

is the extended observability matrix and

$$
\mathcal{C}_s = \begin{pmatrix} B & AB & \ldots & A^{s-1}B \end{pmatrix}
$$

the extended controllability matrix.

For $k = 0$, the factorization is then:

$$
\begin{aligned}
H_{r,s}^{(0)} &= \mathcal{O}_r \cdot \mathcal{C}_s \\
&= U \Sigma V^T \\
&= U \Sigma^{1/2} \Sigma^{1/2} V^T
\end{aligned}
$$

Where the second inequality is obtained through singular value decomposition. We then have:

$$
\begin{aligned}
\mathcal{O}_r &= U \cdot \Sigma^{1/2} \\
\Rightarrow \mathcal{O}_r^\dagger &= \Sigma^{-1/2} U^T
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{C}_s &= \Sigma^{1/2} \cdot V^T \\
\Rightarrow \mathcal{C}_s^\dagger &= V^T \Sigma^{-1/2}
\end{aligned}
$$

The dagger sign on e.g. $\mathscr{O}_r^\dagger$ denotes the pseudo inverse.

The state space matrices are then, inserting the expressions for the pseudoinverse for the extended observability and controllability in for example $A = \mathscr{O}_r^\dagger H_{r,s}^{(1)} \mathscr{C}_s^\dagger$ and using the expressions above:

$$\begin{aligned}
\hat{A}_n &= \mathscr{O}_r^\dagger H_{r,s}^{(1)} \mathscr{C}_s^\dagger \\
&= \Sigma_n^{-1/2} U_n^T H_{r,s}^{(1)} V_n \Sigma_n^{-1/2} \\
\hat{B}_n &= \mathscr{C}_s \cdot \begin{bmatrix} I_{m\times m} & 0_{m\times(s-1)m} \end{bmatrix}^T \\
\hat{C}_n &= \left[\begin{pmatrix} I_{p\times p} \\ 0^T_{p\times(s-1)p} \end{pmatrix} \right]^T \mathscr{O}_r
\end{aligned}$$