Solution for the exam in the course “Modeling and Learning for Dynamical System" (TSRT92) 2022-01-07

1. (a) BJ-model is more flexible than ARX-model, because the disturbance-model is modeled separately. With the ARX model, the disturbance-model gets the same poles as the system dynamics. This often means that the number of orders becomes unnecessarily high to be able to describe the disturbance-model and that the ARX model is much simpler to estimate because it is a linear regression (solve a system of equations), which should be compared with numerical search of (local) minimum for the BJ model.

   (b) The solution most often involves parts with slow changes combined with parts with fast changes. It is therefore inappropriate to use the same step length.

   (c) Modelica uses object-oriented model building. This means that you are not sure to get a state-space-model if you connect two systems which in themselves are state-space-models.

   (d) Validation of a model is always necessary for being able to judge whether it can describe the system properties. For cross-validation, another data is used for validation than is used in the model estimation.

   Using the same data for both estimation and validation means that a higher order model always leads to a better curve fit, even if the correct order was passed. The additional parameters are in this case used to describe the specific disturbance signal.

   Using cross-validation, a model with too high order is not favored because it is another disturbance signal than the one that was used for estimating the model.

2. (a) The input signal should have its main frequency content in the frequency band that you want the model to have a good fit. It is a linear system we are going to simulate, so we can choose a binary signal that alternates between -1 and 1. Because we are interested in a good adaptation for low frequencies, the frequency content of the signal is determined by the passband 0-0.1 of the Nyquist frequency, i. e. the range 0 - 0.05 Hz.

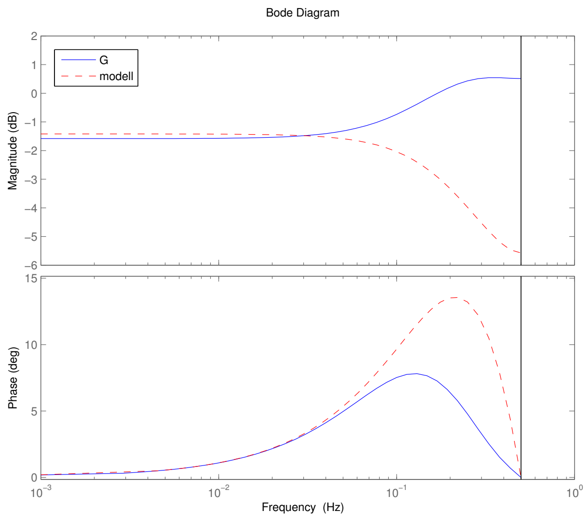
```
G = tf([1, -0.4], [1, -0.3, 0.02], 1, 'variable', 'z^-1')
N = 10000;
u = idinput(N,'rgs',[0 0.1]);
t = 0:T:(N-1)*T;
y = lsim(G, u, t); % simulerar systemet
```

   (b)
```
data = iddata(y, u, T);
m = arx(data, [1 1 1]);
```

   Identification with ARX model structure of order 1 gives $a = 4.262 \pm 0.01014$ och $b = -2.77 \pm 0.008552$ The figure shows that the fit is good for low frequencies.

   (c) Change the properties of the input signal to higher frequencies, e. g. using the following code

```
u = idinput([N 1],'rgs',[0.8 1]);
```



that focuses on frequencies $0.4 - 0.5\,\mathrm{Hz}$.

3. (a) In this case, the parameters are gained from solving the equations

$$G(e^{i\omega_1}, \theta) = G_0(e^{i\omega_1})$$
$$G(e^{i\omega_2}, \theta) = G_0(e^{i\omega_2})$$

where $\theta = (b_1, b_2, f_1, f_2)$ and

$$G(e^{i\omega}, \theta) = \frac{b_1 e^{-i\omega} + b_2 e^{-2i\omega}}{1 + f_1 e^{-i\omega} + f_2 e^{-2i\omega}}$$

This is a system with four parameters and four equations that is uniquely solvable.

(b) It holds that the parameters converge

$$\theta^* = \arg\min_{\theta} \int_{-\pi}^{\pi} G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 \frac{\Phi_u(\omega)}{|H_*(e^{i\omega})|^2} d\omega \tag{1}$$

where $\Phi_u(\omega)$ is the spectrum of the input signal. (See eq. (12.87) in the course book.) In this case we have

$$\Phi_u(\omega) = \sum_{k=1}^{2} \delta(\omega - \omega_k) + \delta(\omega + \omega_k) \tag{2}$$

where δ is the Dirac-impulse and $H_*(e^{i\omega_k}) = 1$. This input spectrum gives then that

$$\theta^* = \arg\min_{\theta} \sum_{k=1}^{2} \left(|G_0(e^{i\omega_k}) - G(e^{i\omega_k}, \theta)|^2 + |G_0(e^{-i\omega_k}) - G(e^{-i\omega_k}, \theta)|^2 \right)$$

which is minimized by choosing from (a), i. e. we get perfect match at the given frequencies.

(c) In the case of three cos-terms we get the input spectrum

$$\Phi_u(\omega) = \sum_{k=1}^{3} \delta(\omega - \omega_k) + \delta(\omega + \omega_k)$$

which gives at θ^* is minimized

$$\theta^* = \arg\min_\theta \sum_{k=1}^{3} \left(|G_0(e^{i\omega_k}) - G(e^{i\omega_k}, \theta)|^2 + |G_0(e^{-i\omega_k}) - G(e^{-i\omega_k}, \theta)|^2 \right)$$

Unlike the cases above, we can not be sure of getting a perfect match for the given frequencies because the number of parameters is less than the number of equations.

(d) In this case, the input spectrum from eq. (2) is used and θ^* is given by

$$\theta^* = \arg\min_\theta \sum_{k=1}^{2} \left(\frac{|G_0(e^{i\omega_k}) - G(e^{i\omega_k}, \theta)|^2}{|H_*(e^{i\omega_k})|^2} + \frac{|G_0(e^{-i\omega_k}) - G(e^{-i\omega_k}, \theta)|^2}{|H_*(e^{-i\omega_k})|^2} \right)$$

Even in this case, regardless of H^*, you can select the parameters so that similarity is achieved. That is, in the three cases where the input signal is a sum of two cosine terms the same θ^* is obtained. However, the input signal with three cosine terms may give a different estimate.

4. (a) Let $a_0 = A^T(AA^T)^{-1}y$ and let $a = a_0 + \Delta a$. We then have

$$Xa = Xa_0 + X\Delta a = y + X\Delta a$$

and hence $X\Delta a = 0$ for a to be a solution. Now consider

$$\|a\|_2^2 = \|a_0 + \Delta a\|_2^2 = \|a_0\|_2^2 + \|\Delta a\|_2^2 + 2\Delta a^T a_0$$

where

$$\Delta a^T a_0 = \Delta a^T X^T (XX^T)^{-1} y = 0$$

Hence we see that any a that satisfies the equation has at least the norm of a_0, and therefore a_0 is the minimum norm solution.

(b) We have that

$$\frac{\partial f_k(a)}{\partial a} = 2(y_k - a^T x_k)x_k$$

and hence if we start with $a_0 = 0$, we obtain that the solution we obtained from the gradient algorithm is of the form

$$a = \sum_{k=1}^{N} \alpha_k x_k$$

for some $\alpha_k \in \mathbb{R}$. Equivalently we may write $a = X^T\alpha$, where $\alpha = \begin{bmatrix}\alpha_1 & \cdots & \alpha_N\end{bmatrix}^T$. Together with $Xa = y$ we obtain the following equation for α:

$$XX^T\alpha = y$$

which has a unique solution if X has full row rank. Then it follows that

$$a = X^T \left(XX^T\right)^{-1} y$$

which agrees with the minimum norm solution.