SSY281 Model Predictive Control 2023
PSS 2 - State Estimation and Setpoint Tracking

Rémi Lacombe

February 7th, 2023

Course reminder

Cf PSS notes.

Exercise 1

Unconstrained Tracking Problem

1. For an unconstrained system, show that the following condition is sufficient for feasibility of the target problem for any z_{sp}.

$$\text{rank} \begin{bmatrix} I - A & -B \\ HC & 0 \end{bmatrix} = n + n_c \tag{1}$$

   where n is the number of states, n_c is the number of controlled outputs, and p is the total number of measured outputs. The size of the matrix in (1) is $(n + n_c) \times (n + m)$.
   Find a counter-example to prove that condition (1) is sufficient only, and not necessary.

2. Show that the condition (1) implies that the number of controlled variables without offset is less than or equal to the number of manipulated variables and the number of measurements, i.e. $n_c \leq m$ and $n_c \leq p$.

3. Show that (1) implies the rows of H are independent.

4. Does (1) imply that the rows of C are independent? If so, prove it, if not provide a counter-example.

5. By choosing H, how can one satisfy (1) if one has installed redundant sensors so several rows of C are identical?

Solution:

1. In the unconstrained problem, feasibility depends on the existence of a solution to
$$\begin{bmatrix} I-A & -B \\ HC & 0 \end{bmatrix} \begin{bmatrix} x_s \\ u_s \end{bmatrix} = \begin{bmatrix} 0 \\ z_{sp} \end{bmatrix} \tag{2}$$
This solution exists if the rows of the matrix (1) are independent, which is the rank condition given in the exercise but this condition is not *necessary and sufficient* but only *sufficient* because the right-hand-side in (2) is not arbitrary, but has a vector of zeros on left-hand-side above the controlled variable set-points z_{sp}. For example, if $A = I$ and $B = 0$, a solution can exist ($H = I$, $C = I$) despite violation of rank condition (1).

2. For the rows to be independent, the number of rows $n + n_c$ must be less than or equal to the number of columns $(n + m)$, which gives $n_c \leq m$. Because of the zero in the second row, the rank condition implies also that $\text{rank}(HC) = n_c$. It is useful to recall the following result on the rank of a matrix product here: $\text{rank}(HC) \leq \min(\text{rank}(H), \text{rank}(C))$. Since H is an $n_c \times p$ matrix, we need to have $n_c \leq p$ to meet the rank condition on HC. Although the exercise doesn't ask the question, we can also conclude that $n_c \leq n$, i.e. the number of controlled outputs should be less than or equal to the number of states.

3. The n_c rows of H are independent otherwise we contradict $\text{rank}(HC) = n_c$ from the previous part.

4. The rows of C do not need to be independent to satisfy the given rank condition. A simple counter-example is the case where
$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3}$$
But note that we cannot choose all of the outputs corresponding to linearly dependent rows as controlled variables and satisfy the rank condition. So e.g. $H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ would not work here.

5. We can have redundant sensors and choose the mean of the redundant sensors as part of the controlled outputs, for example. In the previous example, we can imagine that the two first outputs are given by redundant sensors on the first state component. A reasonable choice for H to leverage this information fully would then be $H = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Exercise 2

Redundant sensors
A constant variable x is measured by two different sensors with different accuracy. The system is described by

$$
\begin{aligned}
x(k+1) &= x(k) \\
y(k) &= Cx(k) + e(k)
\end{aligned}
$$

with

$$C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and $e(k)$ is a zero-mean white-noise vector with covariance matrix

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}$$

1. Estimate x as

   $$\hat{x}(k) = a_1 y_1(k) + a_2 y_2(k)$$

   and determine the constants a_1 and a_2 so that the mean value of the estimation error is zero and so that the variance of the estimation error is as small as possible. Compare the variance with that obtained when using only one of the sensors.

2. Compare your solution above with that obtained with the stationary Kalman filter.

Solution:

MMSE (Minimum Mean Square Error) Estimator:

1. The constants a_1 and a_2 are determined using following two conditions

   (a) The mean value of estimation error should be zero.

   (b) The variance of estimation error should be minimized

   Case 1: Both Sensors are used

   The estimator is given by

   $$\hat{x}(k) = a_1 y_1(k) + a_2 y_2(k)$$

   and the estimation error is calculated as

   $$
   \begin{aligned}
   \tilde{x}(k) &= x(k) - \hat{x}(k) = x(k) - a_1 x(k) - a_1 e_1(k) - a_2 x(k) - a_2 e_2(k) \\
   \tilde{x}(k) &= (1 - a_1 - a_2)x(k) - a_1 e_1(k) - a_2 e_2(k)
   \end{aligned}
   $$

Note that $e_1(k)$ and $e_2(k)$ are zero-mean white noises i.e. $E[e_1(k)] = E[e_2(k)] = 0$. According to condition 1,

$$E[\tilde{x}(k)] = (1-a_1-a_2)E[x(k)]-a_1 \underbrace{E[e_1(k)]}_{=0} - a_2 \underbrace{E[e_2(k)]}_{=0} = (1-(a_1+a_2))E[x(k)] = 0, \tag{4}$$

and since $E[x(k)] \neq 0$, it implies

$$a_1 + a_2 = 1.$$

The variance of estimation error is given by[1]

$$\mathrm{Var}(\tilde{x}(k)) = E[\tilde{x}(k)^2] - (\underbrace{E[\tilde{x}(k)]}_{=0})^2 = E[\tilde{x}(k)^2] = E[((1-(a_1+a_2))x(k) - a_1 e_1(k) - a_2 e_2(k))^2]$$

Note that both output noises $e_1(k)$ and $e_2(k)$ are zero-mean white noises with variance $\sigma_{e_1}^2 = 1$ and $\sigma_{e_2}^2 = 9$ respectively. Also note that $e_1(k)$ and $e_2(k)$ are uncorrelated thus $E[e_1(k)e_2(k)] = 0$. Thus expanding above expression and then using this data we get

$$\mathrm{Var}(\tilde{x}(k)) = (1 - a_1 - a_2)^2 E[x(k)^2] + a_1^2 \cdot 1 + a_2^2 \cdot 9$$

since $1 - (a_1 + a_2) = 0$, we get

$$\mathrm{Var}(\tilde{x}(k)) = a_1^2 + 9 \cdot (1 - a_1)^2$$
$$\mathrm{Var}(\tilde{x}(k)) = 10a_1^2 + 9 - 18a_1$$

According to condition 2, constants a_1 and a_2 should be chosen such that we get the minimum value of the variance. Thus, in order to find the optimal parameter values, let us define $V = \mathrm{Var}(\tilde{x}(k))$ as the objective function/criterion to be minimized. Now taking the derivative of V with respect to a_1 we get

$$\frac{\mathrm{d}V}{\mathrm{d}a_1} = 20a_1 - 18 = 0$$

$$\Rightarrow a_1 = \frac{9}{10} \Rightarrow a_2 = \frac{1}{10}$$

The estimator using two measurements is thus given by

$$\hat{x}(k) = \frac{9}{10} y_1(k) + \frac{1}{10} y_2(k)$$

Using a_1 and a_2, the minimum (or optimal) value of variance is given by

$$V^* = \frac{9}{10}$$

[1]Note that the variance of any random variable X, denoted by $\mathrm{Var}(X)$ or σ_X^2 is defined as $\mathrm{Var}(X) = \sigma_X^2 = E[(X - E[X])^2] = E[X^2] - (E[X])^2$ where $E[X^2]$ is the mean power or mean-square-error (MSE) of X and $E[X] = m_X$ is the mean of X. Also recall that for independent random variables X_i:s, $\mathrm{Var}(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n + b) = a_1^2 \mathrm{Var}(X_1) + a_2^2 \mathrm{Var}(X_2) + \cdots + a_n^2 \mathrm{Var}(X_n)$.

Case 2: Only one sensor is used

Let us suppose only one of the sensor 'i' is available for measurement. The estimator in this case is simply given by

$$\hat{x}(k) = a_i y_i(k).$$

The estimation error is given by

$$\begin{aligned}\tilde{x}(k) &= x(k) - \hat{x}(k) = x(k) - a_i x(k) - a_i e_i(k)\\ \tilde{x}(k) &= (1 - a_i)x(k) - a_i e_i(k)\end{aligned}$$

Note that the coefficient a_i should be naturally equal to 1. However, we can also verify this by applying condition 1 according to which

$$E[\tilde{x}(k)] = (1 - a_i)E[x(k)] - a_i \underbrace{E[e_i(k)]}_{=0} = (1 - a_i)E[x(k)] = 0 \qquad (5)$$

Since $E[x(k)] \neq 0$, it implies

$$a_i = 1.$$

The variance of estimation error is given by

$$\mathrm{Var}(\tilde{x}(k)) = E[\tilde{x}(k)^2] - (\underbrace{E[\tilde{x}(k)]}_{=0})^2 = E[\tilde{x}(k)^2] = E[((1-a_i)x(k) - a_i e_i(k))^2]$$

$$\mathrm{Var}(\tilde{x}(k)) = (1 - a_i)^2 E[x(k)^2] + a_i^2 \sigma_i^2$$

Since $a_i = 1$, we get

$$\mathrm{Var}(\tilde{x}(k)) = \sigma_i^2 = V^*$$

which is the minimum possible variance in our estimate that we can get, where $\sigma_i^2 = 1$, if sensor 1 ($i = 1$) is used and $\sigma_i^2 = 9$, if sensor 2 ($i = 2$) is used.

Remark. *Using only measurement y_1 gives variance 1 and if only y_2 is used then the minimum variance is 9. Thus using redundant sensor and combining two measurements gives a better result than using only one of the best measurement.*

Kalman Estimator:

2. Now let us compare the previous estimator with *steady-state* Kalman filter. Assume that a priori estimate of x is zero and the *prediction error covariance* of x is p, i.e.

$$\hat{x}(0|0) = 0 \;\text{ and }\; P(k|k) = P(0|0) = p,$$

and the sensor measurement noise covariance is given by

$$R = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}.$$

Recall that the model has no process noise, therefore, $Q = q = 0$. Also note that we use steady-state Kalman filter so covariance $P = p$ is constant. The steady-state Kalman gain is given by

$$L = PC^T[CPC^T + R]^{-1} = p\begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \frac{1}{10p+9} \cdot \begin{bmatrix} p+9 & -p \\ -p & p+1 \end{bmatrix}$$
$$\Rightarrow L = \frac{p}{10p+9} \cdot \begin{bmatrix} 9 & 1 \end{bmatrix}.$$

Now using gain L, the steady-state Kalman estimator is given by

$$\hat{x}(k|k) = \hat{x}(k|k-1) + L(y(k) - C\hat{x}(k|k-1)) = (1 - LC)\hat{x}(k|k-1) + Ly(k)$$
$$\hat{x}(k|k) = \left(1 - \frac{10p}{10p+9}\right)\hat{x}(k|k-1) + \frac{9p}{10p+9}y_1(k) + \frac{p}{10p+9}y_2(k)$$
$$\hat{x}(k|k) = \frac{9}{10p+9}\hat{x}(k|k-1) + \frac{9p}{10p+9}y_1(k) + \frac{p}{10p+9}y_2(k)$$
$$\hat{x}(k|k) = a_3\hat{x}(k|k-1) + a_1y_1(k) + a_2y_2(k)$$

Note that the Kalman filter estimates the state at current time step by using the optimal combination of model-based state prediction and the sensor readings. The values of weights a_i:s will depend on relative level of process and measurement noises (Q and R).

Relationship between MMSE and Kalman Estimators

Let us evaluate the the Kalman estimator for the case $p \to \infty$. The weights under the limit are given by

$$\lim_{p\to\infty} a_1 = \frac{9}{10}, \quad \lim_{p\to\infty} a_2 = \frac{1}{10}, \quad \lim_{p\to\infty} a_3 = 0.$$

This shows that for a case where p (i.e. when prediction error covariance) is very large then the weights for y_1 and y_2 will be approximately equal to those calculated for the MMSE estimator in part a. In other words, Kalman Estimator is equivalent to MMSE estimator if prediction error covariance is very high. The reason is that if the process model is highly uncertain (i.e. p is very large) then the Kalman Filter will show almost no trust on the prediction $\hat{x}(k|k-1)$. It will instead show high trust on the sensor measurements and will optimally combine two sensors readings, which should naturally lead to the same optimal result obtained by the MMSE estimator as shown in part a.