POLITECNICO DI MILANO





Introduction to the state estimation problem

Marco Lovera

Dipartimento di Scienze e Tecnologie Aerospaziali, Politecnico di Milano

Problem statement

For the LTI stochastic system

$$\dot{x} = Ax + Bu + w, \qquad x(0) = x_0$$
$$y = Cx + Du + v$$

where *v* and *w* are white Gaussian noise processes and x_0 is a Gaussian random variable, we want to define estimators for the state vector *x* on the basis of measurements of the input *u* and of the output *y*.

Problem statement

Interpretation of *v* and *w*:

- *v* represents measurement noise acting on the output of te system;
- *w* is process noise, accounting for disturbances and model uncertainty.

Assuming we have data over the time interval from 0 to *T*, we denote the estimate of the state

$$\widehat{x}(t|T)$$

and formulate three different problems.

Problem statement

- $t > T$: *prediction* problem.
- $t = T$: *filtering* problem.
- $0 < t < T$: *smoothing* problem.

Furthermore, the problem can be formulated in different ways depending on the modelling assumptions:

- CT dynamics and CT measurement process
- CT dynamics and DT measurement process (as in the OE problem)
- DT dynamics and DT measurement process.

We will study all three formulations, in reverse order.

Possible applications

- Generalisation of OE to the case in which process noise is not negligible. In such a case the likelihood function depends also on process noise through the state sequence, which is no longer deterministic.

- Control-oriented state estimation. Many modern control design methods assume that *state feedback* is possible, *i.e.*, that we can measure the state variables (with negligible noise).

  This is usually not possible, hence the need for estimators capable of recostructing the states from measurements.

Possible applications

- Flight test data analysis: smoothing allows to reconstructed un-measurable states using telemetry data for post-flight analyses.

- Model-based fault detection and isolation: filtering and prediction of a plant's states and outputs based on a nominal model can be used to detect faults, *i.e.,* changes in the dynamics of the plant with respect to the nominal one.

A deterministic approach

- Assume that there are no stochastic inputs and that the initial state is known exactly:

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0$$
$$y = Cx + Du.$$

- Then *in principle* one should be able to estimate the state of the system just by simulating the state equation.

- In this respect, output measurements are not needed, as only input measurements would be used!

A deterministic approach

- This elementary idea fails as soon as disturbances come into play:
  - The simulation diverges from the real state
  - There is no way to implement a correction to bring the estimation error down.

A deterministic approach

- An alternative approach making use of output data is the following (illustrated first in the scalar case).

- The plant dynamics is given by:

$$\dot{x} = ax + bu, \qquad x(0) = x_0$$
$$y = cx$$

- The estimator is defined as

$$\dot{\hat{x}} = a\hat{x} + bu + l(y - \hat{y}), \qquad x(0) = \hat{x}_0$$
$$\hat{y} = c\hat{x}$$

- How does this estimator work?

$$\dot{\hat{x}} = a\hat{x} + bu + l(y - \hat{y}), \quad x(0) = \hat{x}_0$$
$$\hat{y} = c\hat{x}$$

- When the error between true and estimated output is small the last term is close to zero and we are basically simulating the plant model.

- On the other hand, when the output estimation error is large, the last term provides a correction term.

Time-domain analysis

- We study the dynamics of the state estimation error, defined as

$$e = x - \widehat{x}.$$

- Differentiating we have

$$\begin{aligned}\dot{e} &= \dot{x} - \dot{\widehat{x}} = \\ &= ax + bu - a\widehat{x} - bu - lc(x - \widehat{x}) = \\ &= (a - lc)(x - \widehat{x}) = \\ &= (a - lc)e.\end{aligned}$$

A deterministic approach

- Therefore, if the gain *l* is chosen such that

$$(a - lc) < 0$$

then the error dynamics

$$\dot{e} = (a - lc)e$$

is asymptotically stable and no matter what the initial error is we have that

$$e \underset{t \to \infty}{\longrightarrow} 0.$$

Frequency-domain analysis.

Assuming that $(a - lc) < 0$ we can study the estimator in the frequency domain.

In particular we can compute the transfer function from the true state to the estimated one.

The estimator

$$\dot{\hat{x}} = a\hat{x} + bu + l(y - \hat{y}), \quad x(0) = \hat{x}_0$$
$$\hat{y} = c\hat{x}$$

can be written as

$$\dot{\hat{x}} = (a - lc)\hat{x} + bu + ly, \quad x(0) = \hat{x}_0$$

and recalling the output equation of the plant we have

$$\dot{\hat{x}} = (a - lc)\hat{x} + bu + lcx, \quad x(0) = \hat{x}_0$$

A deterministic approach

Therefore the transfer function from true to estimated state is given by

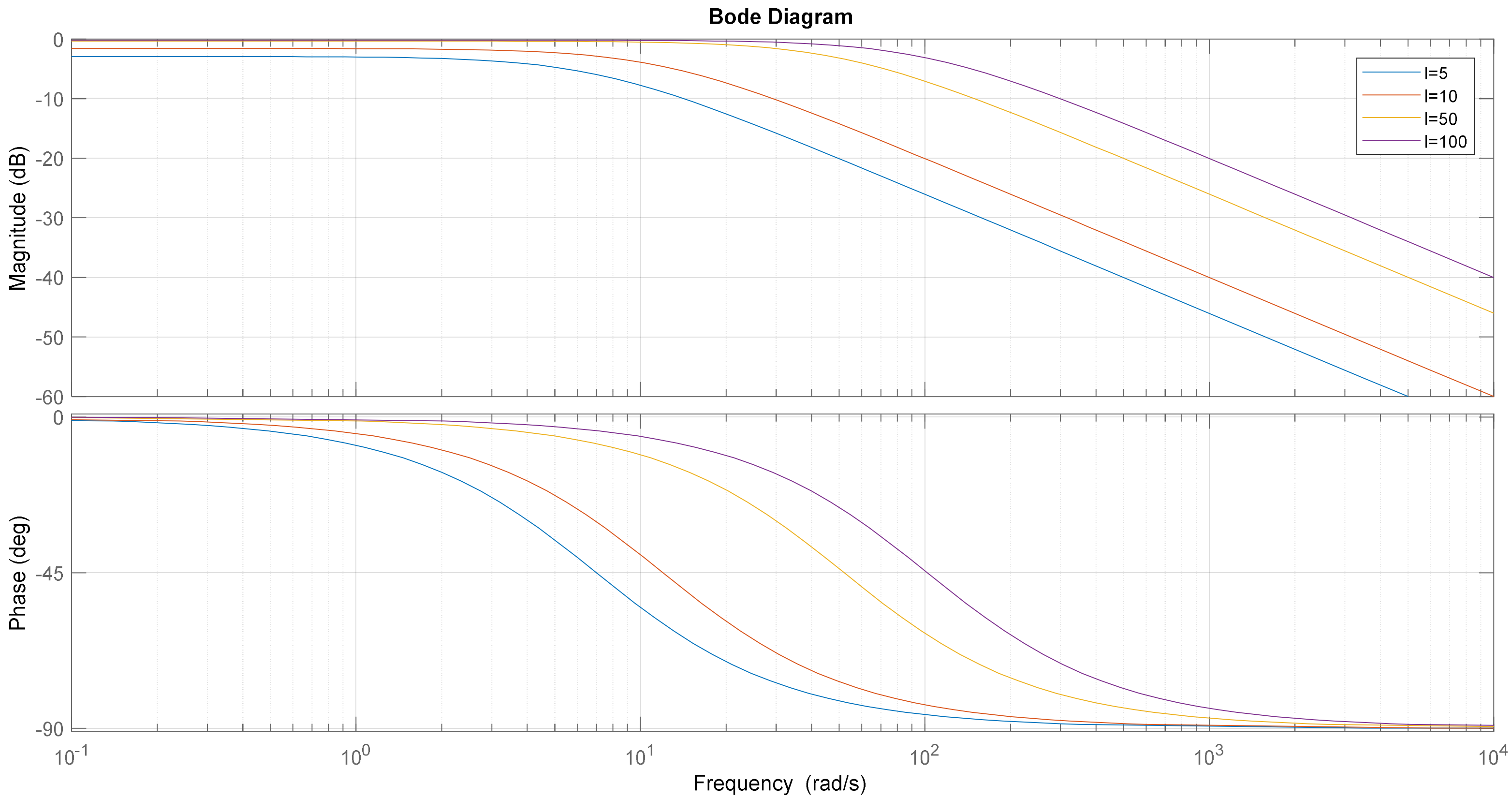
$$G(s) = \frac{lc}{s - (a - lc)}.$$

For example: if *a*=-2, *c*=1 to have $(a - lc) < 0$ the gain must satisfy

$$l > \frac{a}{c} = -2.$$

Let's look at the FRF of the transfer function for increasing values of *l*.

A deterministic approach



Comments:

- For increasing l the gain of the transfer function converges to 1 so at low frequency the steady state error becomes smaller and smaller.
- For increasing l the bandwidth of the filter also increases, so faster and faster components of the true state are estimated correctly.
- Can we define criteria to choose l in an optimal way? For example, taking into account properties of measurement and process noise?

A deterministic approach

- For a system of order *n* with *p* outputs we can proceed exactly in the same way.

- The plant dynamics is given by:

$$\dot{x} = Ax + Bu, \quad x(0) = x_0$$
$$y = Cx$$

- The estimator is defined as

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}), \quad x(0) = \hat{x}_0$$
$$\hat{y} = C\hat{x}$$

Time-domain analysis

- We study the dynamics of the state estimation error, defined as

$$e = x - \hat{x}.$$

- Differentiating we have

$$\begin{aligned}
\dot{e} = \dot{x} - \dot{\hat{x}} = \\
&= Ax + Bu - A\hat{x} - Bu - LC(x - \hat{x}) = \\
&= (A - LC)(x - \hat{x}) = \\
&= (A - LC)e.
\end{aligned}$$

A deterministic approach

- Therefore, if the gain L is chosen such that

$$\lambda_i(A - LC) < 0, \quad i = 1, \dots, n$$

then the error dynamics

$$\dot{e} = (A - LC)e$$

is asymptotically stable and no matter what the initial error is we have that

$$e \underset{t \to \infty}{\to} 0.$$

A deterministic approach

Facts from linear control theory:

- If the (A,C) is completely observable then it is always possible to find a gain L such that $(A\text{-}LC)$ has all eigenvalues with negative real part.

- Furthermore, under the same assumption it is always possible to choose L such that the eigenvalues of $(A\text{-}LC)$ are *assigned* to specified locations in the complex plane.

Therefore, under observability it is possible to assign the dynamics of the estimator, just like in the scalar case.

In the vector case the need for a systematic rule to tune the gain is even stronger.

A deterministic approach

- In systems theory the estimator

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}), \quad x(0) = \hat{x}_0$$
$$\hat{y} = C\hat{x}$$

is known as the Luenberger state observer and its derivation goes back to the 1950s.

- Important observation: the error dynamics DOES NOT depend on the input *u* so state estimation problems can be equivalently studied with reference to the model

$$\dot{x} = Ax, \quad x(0) = x_0$$
$$y = Cx.$$

Problem statement

If we now return to the stochastic formulation of the problem:

$$\dot{x} = Ax + w, \quad x(0) = x_0$$
$$y = Cx + v$$

where v and w are white Gaussian noise processes and x_0 is a Gaussian random variable, the questions we need to answer are the following:

- How do we quantify the performance of the estimator in a stochastic sense?
- Is the Luenberger estimator structure the optimal one for the stochastic problem?
- If so, how do we find the optimal gain?

Problem statement

Key point:

Unlike the estimation problems studied so far in which

- measurements were random but the parameters to be estimated were unknown constants

in this case

- what we are trying to estimate, *i.e.,* the state vector of the system, is also a random variable.

Therefore this problem cannot be handled using the same methods, *i.e.,* ML estimation theory and we need to resort to so-called Bayesian estimation.