Nonlinear observers

design techniques, part 2

Erik Frisk

Dept. of Electrical Engineering, Linköping University

January 30, 2019

LINKÖPING UNIVERSITY

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- High gain observers
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- High gain observers
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

Which problem are we solving?

For a model

$$\dot{x} = f(x)$$
$$y = h(x)$$

a candidate observer is

$$\dot{\hat{x}} = f(\hat{x}) + K(t)(y - h(\hat{x}))$$

The dynamics for the error $e = x - \hat{x}$ is then

$$\dot{e} = f(x) - f(\hat{x}) - K(t)(h(x) - h(\hat{x}))$$

If the functions $f(x) = Ax$ and $h(x) = Cx$ had been linear, we would have

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

and we could control convergence. What to do in general?

The basic problem

there are some basic methods to achieve observer stability, i.e., error dynamics stability. Common examples are:

- Transform the original model to specific forms where design and stability is easier
- Lyapunov design
  assume a Lyapunov function $V(e)$ and choose $K(t)$ such that $\dot{V} < 0$.
- High-gain
  Choose observer feedback gain large enough
- Sliding mode
  Control the error against a surface where convergence is guaranteed.

Reading material

- Sliding mode observers: I like Barbot [19]. They have material ob transformation to a triangular form; possibly an advanced text. Skip any parts that are not accessible to you. Also the concluding proofs can be omitted. The main parts are 4.1, 4.2, 4.4. The paper [10] is also worth a read.
- High gain observers, read chapter in "Nonlinear Systems" by Khalil. [11]
- Survey papers [3,4] good, possibly a little old but gives a good overview.
- Besancon [7] has a few Lyapunov related results
- Moving Horizon Estimation, main paper [24] Rao et.al. Björn also recommends [25, 26], the course lecture [27] and book [28].

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- High gain observers
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

Design via transformation to linear problem

If we can, using a change of coordinates $z = T(x)$ and possibly transformation of the output $w = \Psi(y)$, can transform the original model

$$\begin{aligned} \dot{x} &= f(x) \\ y &= h(x) \end{aligned}$$

to

$$\begin{aligned} \dot{z} &= Az + g(w) \\ w &= Cz \end{aligned}$$

where (A, C) is an observable pair, then the design problem is trivial.

A simple non-linear problem

$$\dot{x} = Ax + g(y), \quad y = Cx$$

where we have the observer

$$\dot{\hat{x}} = A\hat{x} + g(y) + K(y - C\hat{x})$$

which can easily be made stable if (A, C) is an observable pair since the error dynamics is given by

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

This indicates that this is not always possible; but possibly a little more often than you think.

Globally linearising methods

The basic idea, in a simple case, is to find a state transformation $z = T(x)$ such that you get a non-linear observable canonical form, estimate z and then invert $T(x)$ to get your estimate of x.

For a control affine model

$$\dot{x} = f(x) + g(x)u, \quad y = h(x)$$

there are necessary *and sufficient* (!) conditions for the existence of such a $T(x)$.

A small example

A small example that illustrates the principle

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 x_2 \\ y &= x_1 \end{aligned}$$

Use the change of variables

$$z = T(x) = \begin{pmatrix} x_1 \\ -\frac{1}{2}x_1^2 + x_2 \end{pmatrix}, \quad x = T^{-1}(z) = \begin{pmatrix} z_1 \\ \frac{1}{2}z_1^2 + z_2 \end{pmatrix}$$

In the new variables z we get

$$\begin{aligned} \dot{z} &= \frac{\partial T(x)}{\partial x} f(x)\Big|_{x=T(z)} = \begin{pmatrix} 1 & 0 \\ -x_1 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 x_2 \end{pmatrix}\Big|_{x=T(z)} = \begin{pmatrix} z_2 + \frac{1}{2}z_1^2 \\ 0 \end{pmatrix} \\ y &= h(T^{-1}(z)) = z_1 \end{aligned}$$

A small example

In the new variables z we have

$$\begin{aligned} \dot{z} &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} z + \begin{pmatrix} \frac{1}{2}y^2 \\ 0 \end{pmatrix} \\ y &= \begin{pmatrix} 1 & 0 \end{pmatrix} z \end{aligned}$$

where it is straightforward to design a stabilizing observer.

This also utilized something very useful – output injection

Output injection example

Consider the model with three non-linear maps/lookup tables ($V(p,h)$, $g_1(p)$, and $g_2(p)$)

$$
\begin{aligned}
pV(p,h) &= mRT \\
\dot{m} &= u_1 g_1(p) + u_2 g_2(p) \\
y_1 &= p \\
y_2 &= h
\end{aligned}
$$

After substitution of the measurement signals we have

$$
\begin{aligned}
\dot{m} &= u_1 g_1(y_1) + u_2 g_2(y_1) \\
y_1 V(y_1, y_2) &= mRT
\end{aligned}
$$

use the second equation as a measurement equation to estimate the mass m

$$
\begin{aligned}
\dot{\hat{m}} &= u_1 g_1(y_1) + u_2 g_2(y_1) + K(y_1 V(y_1,y_2) - \hat{m}RT) \\
r &= y_1 V(y_1, y_2) - \hat{m}RT
\end{aligned}
$$

Observable canonical form - linear

$$
\dot{x} = \begin{pmatrix} -a_1 & 1 & 0 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \\ -a_{n-1} & 0 & 0 & \dots & 0 & 1 \\ -a_n & 0 & 0 & \dots & 0 & 0 \end{pmatrix} x + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} u
$$

$$
y = x_1
$$

$$
A - KC = \begin{pmatrix} -(a_1+k_1) & 1 & 0 & 0 & \dots & 0 \\ -(a_2+k_2) & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \\ -(a_{n-1}+k_{n-1}) & 0 & 0 & \dots & 0 & 1 \\ -(a_n+k_n) & 0 & 0 & \dots & 0 & 0 \end{pmatrix}
$$

Nonlinear observable canonical form

One, of many, nonlinear counterparts is the input signal-triangular observable form

$$
\dot{x} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} x + \begin{pmatrix} f_1(x_1, u) \\ f_2(x_1, x_2, u) \\ \vdots \\ f_{n-1}(x_1, \dots, x_{n-1}, u) \\ f_n(x_1, \dots, x_n, u) \end{pmatrix}
$$

$$
y = x_1
$$

By construction, this is locally weakly observable (linearize and you will see). For $n = 3$,

$$
A = \begin{pmatrix} \star & 1 & 0 \\ \star & \star & 1 \\ \star & \star & \star \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}
$$

Nonlinear observable canonical form

Observability (but perhaps not a good estimation approach) is direct

$$
\begin{aligned}
x_1 &= y \\
x_2 &= \dot{x}_1 - f(x_1) = \dot{y} - f(y) \\
x_3 &= \dot{x}_2 - f(x_1, x_2) = \ddot{y} - f(y, \dot{y} - f(y)) \\
x_4 &= \dots
\end{aligned}
$$

Globally linearising methods

- Conditions guaranteeing a solution is, not so surprising, rather strict.
- The conditions quickly becomes technical and "messy" for MIMO but for SISO they are a bit more appetising. See Torkel Glad's text or the book by Isidori for examples.
- can be generalized to more general cases, for example by also allowing output transformations.
- My experience: Nice approach in theory but rarely applicable in practice
- Necessary and sufficient conditions for, e.g., control-affine systems with smooth non-linearities

$$\dot{x} = f(x) + g(x)u$$
$$y = h(x)$$

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- High gain observers
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

Direct Lyapunov design

For a linear system

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

you can assume an observer

$$\dot{\hat{x}} = A\hat{x} + Bu + K(y - C\hat{x})$$

and then the error dynamics is

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

and then choose K such that $A - KC$ is a stable matrix.

Lyapunov, cont.

Instead, assume a quadratic Lyapunov function

$$V(e) = e^T Pe, \quad P > 0$$

Differentiation gives

$$\dot{V} = e^T(P(A - KC) + (A - KC)^T P)e$$

Choose K such that $\dot{V} < 0$, then asymptotic stability is guaranteed, i.e., that

$$\lim_{t\to\infty} e(t) = 0$$

Lyapunov design

The above approach is directly applicable to non-linear systems

$$\dot{x} = f(x, u)$$
$$y = h(x)$$

with an observer

$$\dot{\hat{x}} = f(\hat{x}, u) + K(y - h(\hat{x}))$$

where the error dynamics is described by

$$\dot{e} = f(x, u) - f(\hat{x}, u) + K(h(x) - h(\hat{x}))$$

Here you can also see the fundamental problems with the non-linearities.

Non-linear Lyapunov

It is an art to design V such that it can be used to prove stability and general (simple) methods are note possible. In general, diagonal Lyapunov functions do not work for interconnected systems.

An obviously stable system

$$\dot{e}_1 = -e_1 + e_2$$
$$\dot{e}_2 = -e_2$$

with a diagonal candidate Lyapunov function $V(e) = e_1^2 + \gamma e_2^2$ does not work but with a mixed term $e_1 e_2$ it works. This is due to that $V(t)$ can go to 0 even if $\dot{V}(t)$ is not negative for all t.

Thau's metod

Looking at special cases, there are general results like:

$$\dot{x} = Ax + f(x), \quad y = Cx$$

and $f(x)$ is locally Lipschitz, i.e., $\|f(x_1) - f(x_2)\| < L\|x_1 - x_2\|$ and there is a solution H, K, P with $P > 0$ for the equation

$$Q(A - KC) + (A - KC)^T Q + 2P = 0$$

Then, with a quadratic Lyapunov function $V(e) = e^T Q e$, it is possible to show stability, i.e., $\dot{V} < 0$, if

$$\frac{\lambda_{min}(P)}{\lambda_{max}(Q)} > L$$

Theorem 1.2 in Besancon

A model

$$\dot{x} = A(u)x$$
$$y = Cx$$

with an observer

$$\dot{\hat{x}} = A(u)\hat{x} + S^{-1}C^T(y - C\hat{x})$$
$$\dot{S} = -SA(u) - A^T(u)S + 2C^T C - \theta S$$

Where does this come from?

Use a candidate Lyapunov function $V(e) = e^T S e$ you get

$$\dot{V} = e^T(\dot{S} + SA(u) + A(u)^T - 2C^T C)e = -\theta e^T S e$$

which with $\theta > 0$ and S positive definite we have $\dot{V} < 0$ which gives convergence (note certain similarities with the Kalman filter equations).

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- *High gain observers*
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

High gain observers

Basic idea

Choose observer gain large enough to dominate all non-linear effects in the error dynamics.

To illustrate, consider the model

$$\begin{aligned} x^{(n)} &= f_n(x, \dot{x}, \dots, x^{(n-1)}) \\ y &= x_1 \end{aligned}$$

i.e.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_n &= f_n(x) \\ y &= x_1 \end{aligned}$$

which is a special case of the non-linear observability form.

A second order system

For a second order system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= f_2(x_1, x_2) \\ y &= x_1 \end{aligned}$$

you can consider an observer

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 + K_1(y - \hat{x}_1) \\ \dot{\hat{x}}_2 &= f_2(\hat{x}_1, \hat{x}_2) + K_2(y - \hat{x}_1) \end{aligned}$$

with the error dynamics

$$\dot{e} = \begin{pmatrix} -K_1 & 1 \\ -K_2 & 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \delta(x, e)$$

where $\delta(x, e) = f_2(x) - f_2(\hat{x}) = f_2(x) - f_2(x - e)$.

Example, cont.

In the error dynamics

$$\dot{e} = \begin{pmatrix} -K_1 & 1 \\ -K_2 & 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \delta(x, e)$$

If it weren't for δ we would be finished, K_1 and K_2 are the coefficients in the characteristic polynomial and the poles can be placed arbitrarily.

The influence from δ can be seen in the transfer function from δ to e

$$G(s) = \begin{pmatrix} s + K_1 & -1 \\ K_2 & s \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{s^2 + K_1 s + K_2} \begin{pmatrix} 1 \\ s + K_1 \end{pmatrix}$$

If we could make this 0 we would be happy. That is unfortunately not possible, the idea behind high gain observers is to make this transfer function small.

Example, cont.

Choose $K_2 \gg K_1 \gg 1$, for example as

$$K_1 = \frac{\alpha_1}{\epsilon},\ K_2 = \frac{\alpha_2}{\epsilon^2}$$

then we get

$$G(s) = \frac{\epsilon}{(\epsilon s)^2 + \alpha_1(\epsilon s) + \alpha_2} \begin{pmatrix} \epsilon \\ (\epsilon s) + \alpha_1 \end{pmatrix}$$

which gives that

$$\lim_{\epsilon \to 0} G(s) = 0$$

Can we just choose ϵ to be very small and be happy?

Example, cont.

The error dynamics, in the variables $\eta_1 = e_1/\epsilon$, $\eta_2 = e_2$ can be written as

$$\epsilon\dot{\eta} = \begin{pmatrix} -\alpha_1 & 1 \\ -\alpha_2 & 0 \end{pmatrix} \eta + \epsilon \begin{pmatrix} 0 \\ 1 \end{pmatrix} \delta$$

Solutions will include terms like

$$(1/\epsilon)e^{-at/\epsilon}$$

which tends to impulses as $\epsilon \to 0$ (peaking phenomenon).

With state-feedback from states estimated with a high gain observer, often saturation of the control signal is introduced to minimize the effect of peaking.

Noise

In the observer

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 + K_1(y - \hat{x}_1) \\ \dot{\hat{x}}_2 &= f_2(\hat{x}_1, \hat{x}_2) + K_2(y - \hat{x}_1) \end{aligned}$$

we have that

$$\hat{x} = \begin{pmatrix} s + K_1 & -1 \\ K_2 & s \end{pmatrix}^{-1} \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \delta + \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} y \right)$$

With a high gain observer, the influence of δ is small and the transfer function from y to $\hat{x}$ is

$$G_{\hat{x}y}(s) = \begin{pmatrix} s + K_1 & -1 \\ K_2 & s \end{pmatrix}^{-1} \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} = \dots = \frac{\alpha_2}{(\epsilon s)^2 + \alpha_1(\epsilon s) + \alpha_2} \begin{pmatrix} 1 + \epsilon s \frac{\alpha_1}{\alpha_2} \\ s \end{pmatrix}$$

and

$$\lim_{\epsilon \to 0} G_{\hat{x}y}(s) = \begin{pmatrix} 1 \\ s \end{pmatrix}$$

Summary/comments

- Choose sufficiently high observer gains to dominate the non-linear effects
- Typically described for non-linear canonical forms
- Introduces peaks in estimates
- Noise sensitive
- There are some robustness properties for feedback controllers based on high gain observer estimates.

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- High gain observers
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

Sliding mode, introductory example

For a system, where x_1 is measured

$$\dot{x} = f(x), \quad y = x_1$$

a sliding mode observer typically looks like

$$\dot{\hat{x}} = f(\hat{x}) + K\text{sgn}\,(y - \hat{x}_1)$$

i.e., it has a switching term in the feedback.

Has been shown to have good noise and robustness properties.

Sliding mode controller - basic principle

(very) short summary of the basic principle:

1. Define a surface such that as long as we are on the surface, all states will converge
2. Use a control law to steer against this surface

A multi variable control problem has been reformulated into a scalar control problem.

A differential equation that appears

Stability for

$$\dot{x} = -k\,\text{sgn}(x)$$

is immediate and the function $V(x) = 1/2x^2$ is a Lyapunov function since

$$\frac{1}{2}\frac{d}{dt}x^2 = -kx\,\text{sgn}(x) = -k|x|$$

The solution can be derived as

$$x(t) = \begin{cases} (|x_0| - k\,t)\text{sgn}x_0 & 0 \leq t < \frac{|x_0|}{k} \\ 0 & t \geq \frac{|x_0|}{k} \end{cases}$$

An example

$$\dot{x}_1 = x_2 \qquad \dot{\hat{x}}_1 = \hat{x}_2 + K_1\text{sgn}(y - \hat{x}_1)$$
$$\dot{x}_2 = f(x_1, x_2) \qquad \dot{\hat{x}}_2 = f(y_1, \hat{x}_2) + K_2\text{sgn}(y - \hat{x}_1)$$
$$y = x_1$$

The error dynamics is then

$$\dot{e}_1 = e_2 - K_1\text{sgn}(e_1)$$
$$\dot{e}_2 = f(x_1, x_2) - f(x_1, \hat{x}_2) - K_2\text{sgn}(e_1)$$

Example cont.

Define the surface $S = \{e | e_1 = 0\}$. Choose $K_1 > |e_2|$, then the surface is attractive

$$\frac{d}{dt}\frac{1}{2}e_1^2 = e_1 e_2 - K_1 e_1 \text{sgn}(e_1) < -c|e_1|$$

This gives that $e_1 \to 0$ in finite time, i.e., $e_1(t) = 0,\ t > t_1$. This is referred to as the *sliding condition*

On the surface it holds that the error dynamics for e_2 fulfills

$$\dot{e}_2 = f(x_1, x_2) - f(x_1, \hat{x}_2) - K_2\text{sgn}(e_1) =$$
$$= \Delta f - K_2/K_1 K_1 \text{sgn}(e_1) = \Delta f - K_2/K_1 e_2$$

Simply put, as long as K_2/K_1 is sufficiently large then $e_2 \to 0$ as long as we are on the surface.

The estimation error is "killed" in finite time, one by one.

Design methodology for triangular systems

Consider a model

$$\begin{aligned}
\dot{x}_1 &= x_2 + f_1(x_1, u)\\
\dot{x}_2 &= x_3 + f_2(x_1, x_2, u)\\
&\vdots\\
\dot{x}_{n-1} &= x_n + f_{n-1}(x_1, \dots, x_{n-1}, u)\\
\dot{x}_n &= f_n(x_1, \dots, x_n, u)\\
y &= x_1
\end{aligned}$$

If the system is not in this form, it can sometimes be transformed into this form. Barot et.al. gives some, not easily interpreted, conditions for when this is possible. Will not be covered here, but is similar to the conditions for exact linearization.

Design methodology for triangular systems

The observer is then

$$\begin{aligned}
\dot{\hat{x}}_1 &= \hat{x}_2 + f_1(\hat{x}_1, u) + K_1\,\text{sgn}(x_1 - \hat{x}_1)\\
\dot{\hat{x}}_2 &= \hat{x}_3 + f_2(\hat{x}_1, \hat{x}_2, u) + K_2\,\text{sgn}_1(\tilde{x}_2 - \hat{x}_2)\\
&\vdots\\
\dot{\hat{x}}_{n-1} &= \hat{x}_n + f_{n-1}(\hat{x}_1, \dots, \hat{x}_{n-1}, u) + K_{n-1}\,\text{sgn}_{n-2}(\tilde{x}_{n-1} - \hat{x}_{n-1})\\
\dot{\hat{x}}_n &= f_n(\hat{x}_1, \dots, \hat{x}_n, u) + K_n\,\text{sgn}_{n-1}(\tilde{x}_n - \hat{x}_n)
\end{aligned}$$

where

$$\begin{aligned}
\tilde{x}_2 &= \hat{x}_2 + K_1\,\text{sgn}_1(x_1 - \hat{x}_1)\\
\tilde{x}_3 &= \hat{x}_3 + K_2\,\text{sgn}_2(x_2 - \hat{x}_2)\\
&\dots\\
\tilde{x}_n &= \hat{x}_n + K_n\,\text{sgn}_{n-1}(x_{n-1} - \hat{x}_{n-1})
\end{aligned}$$

$\text{sgn}_i \approx \text{sgn}$ but with anti-peak and LP. See Barbot et.al. for more information.

Properties of a sliding mode observer

- due to noise, you can of course not maintain perfect surface control, but the error dynamics reacts controlled to noise and you are kept close to the surface.
- The (claimed) main advantage compared to, e.g., EKF is robustness properties in function f.
- Many things to say about implementation, e.g., the sgn function is often replaced with a fast saturation function.
- Systems in MIMO companion-form can be handled in the exact same way.

Outline

- Introduction
- Design via transformation to linear problem
- Lyapunov design
- High gain observers
- Sliding mode observer
- Moving Horizon Estimation (MHE)
- Summary

Moving Horizon Estimation

Main idea is to formulate the estimation problem as a non-linear optimization problem utilizing techniques from optimal control.

- Main advantage same as in optimal control; direct to handle constraints
- Reading material
  Rao, Christopher V., James B. Rawlings, and David Q. Mayne. "*Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations.*" IEEE transactions on automatic control 48.2 (2003): 246-258.

$$\begin{aligned} x_{k+1} &= f_k(x_k, w_k) \\ y_k &= h_k(x_k) + v_k \end{aligned}$$

where

$$x_k \in \mathcal{X}_k, w_k \in \mathcal{W}_k, v_k \in \mathcal{V}_k$$

MHE - basic idea

Basic idea is straightforward – reformulate the estimation optimization as an optimization problem

$$\begin{aligned} \min_{x_0, \{w_k\}_{k=0}^{T-1}} \quad & \sum_{k=0}^{T-1} L_k(w_k, v_k) + \Gamma(x_0) \\ s.t. \quad & x_{k+1} = f_k(x_k, w_k) \\ & y_k = h_k(x_k) + v_k \end{aligned}$$

Note that v_k is completely determined by observations y_k, and the optimization variables x_0 and w_k as

$$v_k = y_k - h_k(x_k)$$

MHE - basic idea

$$\min_{x_0,\{w_k\}_{k=0}^{T-1}} \sum_{k=0}^{T-1} L_k(w_k, v_k) + \Gamma(x_0)$$
$$s.t. \quad x_{k+1} = f_k(x_k, w_k)$$
$$y_k = h_k(x_k) + v_k$$

Loss function and initial value loss, for example

$$L(w_k, v_k) = w^T Q_k^{-1} w + v^T R_k^{-1} v,$$
$$\Gamma(x) = (x - x_0)^T \Pi^{-1} (x - x_0)$$

Main problem

Main problem with the above formulation is that, unless there is linear dynamics and quadratic losses (then Kalman Filter), the computational burden increases with new sample and it does not scale well.

MHE

To make the optimization problem tractable; employ a moving horizon approximation. The loss-function can be written as

$$\sum_{k=0}^{T-1} L_k(w_k, v_k) + \Gamma(x_0) = \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \sum_{k=0}^{T-N-1} L_k(w_k, v_k) + \Gamma(x_0)$$

Due to the state-space formulation, the first term only depends on

$$x_{T-N} \text{ and } w_k, k = T-N, \dots, T-1$$

Denote the cost-to-come

$$\mathcal{Z}_\tau(z) = \min \sum_{k=0}^{T-1} L_k(w_k, v_k) + \Gamma(x_0)$$
$$s.t. \quad x(\tau) = z$$

MHE

With the cost-to-come, the moving horizon formulation can be written as

$$\min_{z,\{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} L_k(w_k, v_k) + \mathcal{Z}(z)$$
$$s.t. \quad x_{k+1} = f_k(x_k, w_k)$$
$$y_k = h_k(x_k) + v_k$$

Now, the only problem is what to do with the unknown cost-to-come function

$$\mathcal{Z}(z)$$

Exact expression for cost-to-come infeasible (unless the whole optimization problem is feasible) so we have to accept approximations.

MHE - Cost-to-come approximation

One strategy proposed in Rao et.al. is to use an Extended Kalman Filter, linearized around the state estimates, to approximate the cost-to-come. For a linearized model

$$f_k(x, w) = A_k x + G_k w, \quad h_k(x) = C_k x$$

with

$$L_k(w, v) = w^T Q_k^{-1} w + v^T R_k^{-1} v, \quad \Gamma(x) = (x - \hat{x}_0)^T \Pi^{-1} (x - \hat{x}_0)$$

the arrival cost at time j is

$$\mathcal{Z}_j(z) = (z - \hat{x}_j)^T \Pi_j^{-1} (z - \hat{x}_j) + \Phi_j^*$$

where

$$\Pi_{j+1} = G_j Q_j G_j^T + A_j \Pi_j A_j^T - A_j \Pi_j C_j^T (R_j + C_j \Pi_j C_j^T)^{-1} C_j \Pi_j A_j^T$$

Outline

- *Introduction*
- *Design via transformation to linear problem*
- *Lyapunov design*
- *High gain observers*
- *Sliding mode observer*
- *Moving Horizon Estimation (MHE)*
- *Summary*

Summary

- High-gain
- Sliding-mode
- Lyapunov
- Transformation
- Moving Horizon Estimates
- Some approaches you can see in the literature. Good complements to the more common EKF (with relatives).