New Methods for State Estimation

PAULO OLIVEIRA

DEEC/IST and ISR

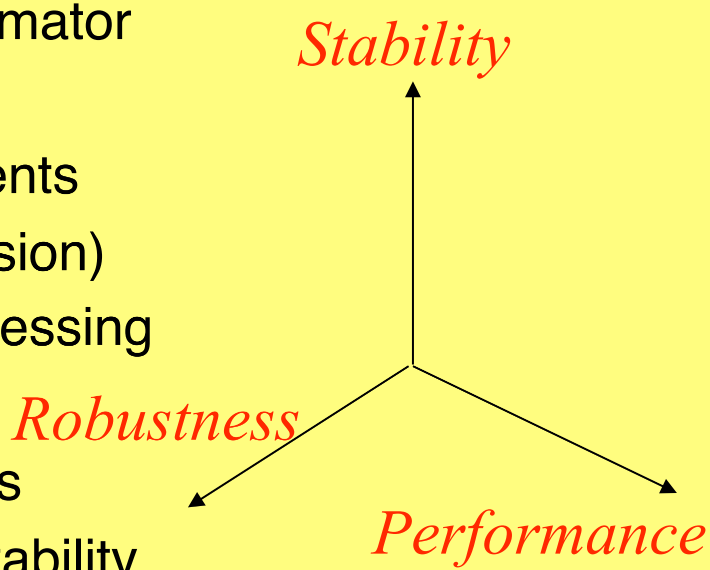
Last Revised: December 14, 2007

Ref. No. NLO#2

Key Challenges in Estimation

Characteristics of the envisioned Estimator

- Reduced computational requirements
- Causal (to be used during the mission)
- Possible to be refined in post-processing



In the linear case, all relevant features are obtained together: exponential stability, optimal performance and robustness (gain and phase margins).

In the nonlinear case no optimal common solution is available.

(e.g. EKF is the performance tentative solution).

Theme

- Stochastic H_2 filtering, prediction, and smoothing problems are only optimal for linear time-varying systems under Gaussian disturbance assumptions with known power spectral densities
- H_∞ allows to lift the noise assumptions for LTV systems
- Real world systems are nonlinear!
- In general, EKF does not guarantee stability, performance, nor robustness
- Nonlinear observers can outperform linear or linearized versions of observers (EKF / SOF), both for structured and unstructured disturbances [1, 2]

Exponential Observers for Linear Systems

Consider the linear system

$$\Sigma_{\mathcal{L}} : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \qquad \begin{array}{l} x(t) \in R^n \\ y(t) \in R^p \\ u(t) \in R^m \end{array}$$

where the pair (A, C) is observable.

The Luenberger observer, in a deterministic setup, is given by

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

Exponential stability can be proven resorting to the Lyapunov equation

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

That is, for a positive definite matrix *Q* there exists a unique positive definite *P*, such that the above equation is verified.

Observers for Nonlinear Systems

Consider the class of affine nonlinear systems

$$\Sigma_{\mathcal{N}} : \begin{cases} \dot{x}(t) = f(x) + g(x)u(t) \\ y(t) = h(x) \end{cases} \qquad \begin{array}{l} x(t) \in R^n \\ y(t) \in R^p \\ u(t) \in R^m \end{array}$$

where $f(.), g(.)$, and $h(.)$ *are known.*

The suggested Luenberger-like nonlinear observer would be

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

What fails in the stability proof
for this nonlinear observer?

Lipschitz Nonlinear Observers

PAULO OLIVEIRA

DEEC/IST and ISR

Last Revised: December 14, 2007
Ref. No. NLO#2

Thau's Observers @ 1973

$$\Sigma_{\mathcal{G}} : \begin{cases} \dot{x}(t) = Ax(t) + f(x(t)) \\ y(t) = Cx(t) \end{cases} \qquad \begin{array}{l} x(t) \in R^n \\ y(t) \in R^p \end{array} \qquad (1)$$

$A, C,$ and $f(.)$ are known; the pair (A, C) is observable, u is a deterministic input, and $f(.)$ is a Lipschitz time-invariant function, i.e.

$$\| f(x(t)) - f(\hat{x}(t)) \| \leq L \| x(t) - \hat{x}(t) \|$$

Proposed observer (motivated by Luenberger's and Kalman's work)

$$\dot{\hat{x}}(t) = A\hat{x}(t) + f(\hat{x}(t)) + K(y(t) - C\hat{x}(t)) \qquad (2)$$

Thau's Observers @ 1973

Proposition : Given the observability condition for systems of class (1) for any positive definite matrix Q, there exists a unique positive definite matrix P, such that the Lyapunov equation

$$A_0^T P + P A_0 = -2Q, \qquad (3)$$

is verified, where $A_0 = A - KC$.

Main result:

Theorem[2] : For the class of systems (1), if the gain K is selected such that it can make the solution of the Lyapunov equation to satisfy

$$\frac{\lambda_{\min}(Q)}{\|P\|} > L \qquad (4)$$

for all x then the Thau observer (2) is assymptotically stable.

Elements of Proof (I)

Defining $\tilde{x}(t) = x(t) - \hat{x}(t)$, and for the Lyapunov candidate function $V(t) = \tilde{x}(t)^T P \tilde{x}(t)$, where $P > 0$ is a symmetric constant matrix, lets apply Lyapunov's second method

$$\dot{V}(t) = \frac{d}{dt}\left[\tilde{x}(t)^T P \tilde{x}(t)\right] = \dot{\tilde{x}}(t)^T P \tilde{x}(t) + \tilde{x}(t)^T P \dot{\tilde{x}}(t)$$

The error dynamics is given by

$$\dot{\tilde{x}}(t) = (A - KC)\tilde{x} + f(x(t)) - f(\hat{x}(t)) = A_0 \tilde{x} + f(x(t)) - f(\hat{x}(t)),$$

and

$$\dot{V}(t) = \tilde{x}(t)^T \left(A_0^T P + P A_0\right)\tilde{x}(t) + 2\tilde{x}(t)^T P\left[f(x(t)) - f(\hat{x}(t))\right]$$

Elements of Proof (II)

From the Lyapunov equation (3) one can write

$$\dot{V}(t) = -2\tilde{x}(t)^T Q\tilde{x}(t) + 2\tilde{x}(t)^T P[f(x(t)) - f(\hat{x}(t))]$$

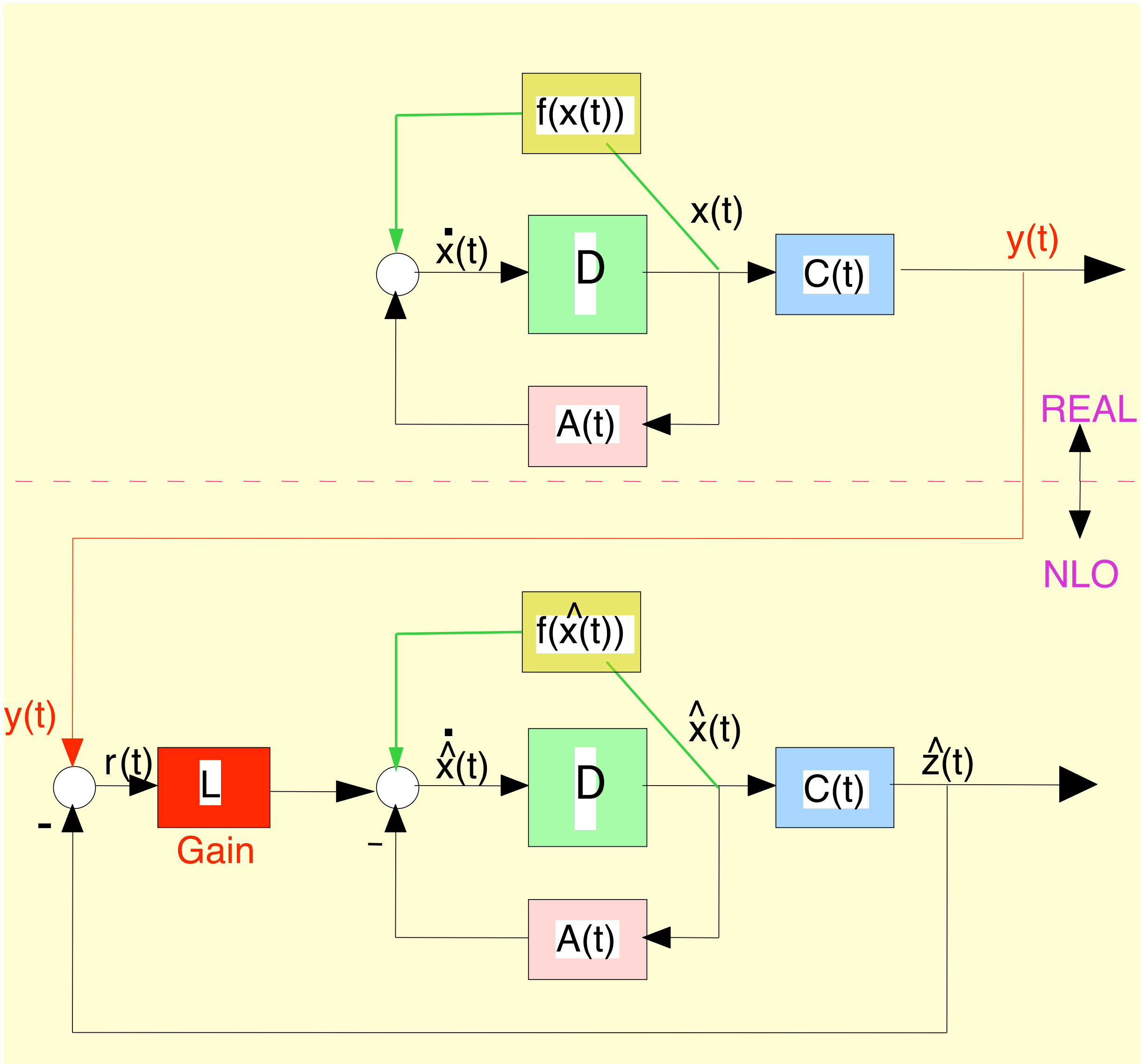
From the Lipschitz condition

$$\begin{aligned}\dot{V}(t) &\leq -2\tilde{x}(t)^T Q\tilde{x}(t) + 2L\|\tilde{x}(t)\|\|P\|\|\tilde{x}(t)\| \\ &\leq -2\lambda_{\min}(Q)\|\tilde{x}(t)\|^2 + 2L\|\tilde{x}(t)\|\|P\|\|\tilde{x}(t)\| \\ &\leq -2[\lambda_{\min}(Q) - L\|P\|]\|\tilde{x}(t)\|^2\end{aligned}$$

if $\lambda_{\min}(Q) > L\|P\|$ then $\dot{V}(t) < 0$, i.e. it is enough that (4) be verified.

The choice of K impacts indirectely on P and constitutes a trial and error process. Moreover, it can lead to very conservative results.

Visualization of the Filter



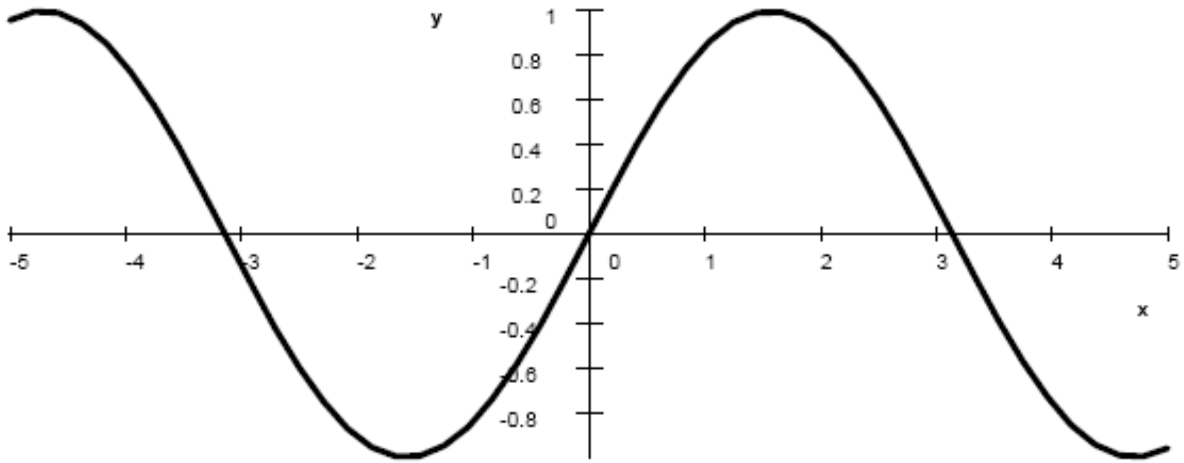
Example with difficulties [4]

Example 1 *nonlinear system*

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u$$

$$y = [1, 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Lipschitz constant for $f(x) = \sin x_1$ *is* $L = 1$. $y = \sin x$



Example with difficulties [4]

The minimum occurs when Q=I.

$$\frac{1}{\|P\|} > 1$$

$$\|P\|^2 = \sum_{i,j=|}^{n} p_{ij}^2 = tr(P^T P)$$

we select $P = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{3} \end{bmatrix}$, $P = P^T > 0$, $\|P\| = 0.68 < 1$, *but* A_0 *no exists. If we select*

The method is not effective, however given *L* and *P*, is of some use to proove stability of the observer.

$\lambda_{\max}(P) = 1.809 > 1$, *the condition is not satisfied. So the method is difficult.*

Exponential Observers @ 1975

Consider the class of nonlinear systems

$$\Sigma_{\mathcal{K}} : \begin{cases} \dot{x}(t) = f(x(t)) & x(t) \in R^n \\ y(t) = h(x(t)) & y(t) \in R^p \end{cases} \qquad (5)$$

where $f(.)$ and $h(.)$ are continuously differentiable.

The proposed observer has the structure

$$\dot{\hat{x}}(t) = f(\hat{x}(t)) + g(y(t), h(\hat{x}(t))) \qquad (6)$$

where $g(.)$ is also continuously differentiable.

An observer should verify in general that

$$g(y(t), h(\hat{x}(t))) = 0 \quad \text{if} \quad h(x(t)) = h(\hat{x}(t)) \qquad (7)$$

however it is impossible to guarantee that $\hat{x}(0) = x(0)$.

Exponential Observers @ 1975

Theorem[3] : Given the nonlinear autonomous systems (5) and (6) and some function $g(.)$ satisfying (7), if there exists a scalar function $V(\tilde{x})$, where $\tilde{x} = x - \hat{x}$ and $\rho > 1$ such that

$$\text{a) } V(\tilde{x}) \geq c\|\tilde{x}\|, \text{ for all } \tilde{x} \in R^n, V(0) = 0$$

$$b) \dot{V}(\tilde{x}) \leq -\lambda V(\tilde{x}), \text{ for all } \tilde{x} \in R^n \text{ and for some } \lambda > 0.$$

then the observer (6) is exponentially stable and

$$\|x(t) - \hat{x}(t)\| \leq k e^{-\lambda t}$$

The generic proof for autonomous systems is an immediate consequence of the Lyapunov second method.

Exponential Observers @ 1975

Main result:

Theorem[3] : Given the nonlinear autonomous system (5) and the proposed observer (6) with $g(y(t), h(\hat{x}(t))) = K(y(t) - h(\hat{x}(t)))$, if there exists an n x m gain matrix K, that, given $Q > 0$ there exits $P > 0$ such that

$$(\nabla f(x) - K\nabla g(x))^T P + P(\nabla f(x) - K\nabla g(x)) = -2Q \qquad (8)$$

From any $\hat{x}(t_0)$, an exponential observer (6) is obtained verifying

$$\|x(t) - \hat{x}(t)\| \leq \left(\frac{q_2}{q_1}\right)^{1/2} e^{-\lambda(t-t_0)} \|x(t_0) - \hat{x}(t_0)\|$$

for all $t > t_0$, where q_1 and q_2 are the smallest and largest eigenvalues of Q, respectively.

Elements of Proof (I)

Defining $\tilde{x}(t) = x(t) - \hat{x}(t)$, the error dynamics is given by

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

For the Lyapunov candidate function $\mathrm{V}(\mathrm{t}) = \tilde{x}(t)^T P \tilde{x}(t)$,

where P > 0 is a symmetric constant matrix (2nd method)

$$\dot{V}(t) = \frac{d}{dt}\tilde{x}(t)^T P \tilde{x}(t) = \dot{\tilde{x}}(t)^T P \tilde{x}(t) + \tilde{x}(t)^T P \dot{\tilde{x}}(t)$$

$$= (f(x(t)) - f(\hat{x}) - K(h(x) - h(\hat{x}))^T P \tilde{x}$$

$$+ \tilde{x}^T P(f(x) - f(\hat{x}) - K(h(x) - h(\hat{x}))$$

* Explicit dependence on *t* will be omitted.

Elements of Proof (II)

Resorting to the fundamental theorem of integral calculus

$$f(x)-f(\hat{x})-K(h(x)-h(\hat{x}))=\int_0^1(\nabla f(w_s)-K\nabla h(w_s))\tilde{x}\,ds$$

for $w_s = sx+(1-s)\hat{x}$. Therefore

$$\dot{V}(t)=\tilde{x}^T\int_0^1(\nabla f(w_s)-K\nabla h(w_s))^T P + P(\nabla f(w_s)-K\nabla h(w_s))ds\,\tilde{x}$$

Using (8) and the fact that $\tilde{x}^T P\tilde{x} > \varepsilon\|\tilde{x}\|^2$ results

$$\dot{V}(t)< -2\varepsilon\tilde{x}^T Q\tilde{x} < -2\frac{q_2}{q_1}\varepsilon V(t), \qquad \text{for } \varepsilon > 0.$$

Example, from [3]

EXAMPLE 1. Consider the following nonlinear system:

$$\dot{x}_1 = x_1\,, \qquad \dot{x}_2 = x_1 - 2x_2 + e^{-x_2}, \qquad y = x_1 + x_2$$

then the gradients of $\mathbf{f}$ and $\mathbf{h}$ are

$$\nabla\mathbf{f} = \begin{bmatrix} 1 & 0 \\ 1 & -2 - e^{-x_2} \end{bmatrix} \quad \text{and} \quad \nabla h = [1 \;\; 1].$$

Let B be a 2×1 constant matrix with elements b_1, b_2 to be determined, then

$$\nabla\mathbf{f} - B\nabla\mathbf{h} = \begin{bmatrix} 1 - b_1 & -b_1 \\ 1 - b_2 & -2 - e^{-x_2} - b_2 \end{bmatrix}.$$

The symmetric part of $\nabla\mathbf{f} - B\nabla h$ is

$$(\nabla\mathbf{f} - B\nabla h)_{sym} = \begin{bmatrix} 1 - b_1 & \frac{1}{2}(1 - b_2 - b_1) \\ \frac{1}{2}(1 - b_2 - b_1) & -2 - e^{-x_2} - b_2 \end{bmatrix}.$$

Thus, if the matrix is selected to be $\left[\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right]$ then

$$(\nabla\mathbf{f} - B\nabla h)_{sym} = \begin{bmatrix} -1 & 0 \\ 0 & -1 - e^{-x_2} \end{bmatrix}. \tag{16}$$

In (16), the two eigenvalues are -1 and $-1 - e^{-x_2}$, so the maximum eigenvalue is -1, i.e.,

$$\mathbf{w}^T(\nabla\mathbf{f} - B\nabla h)\mathbf{w} = \mathbf{w}^T(\nabla\mathbf{f} - B\nabla h)_{sym}\mathbf{w} \leqslant (-1)\cdot\|\mathbf{w}\|^2.$$

Now by Theorem 2 we have that

$$\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z}) + \begin{bmatrix} 2 \\ -1 \end{bmatrix}(\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{z}))$$

or

$$\dot{z}_1 = z_1 + 2[y - (z_1 + z_2)]$$
$$\dot{z}_2 = z_1 - 2z_2 + e^{-z_2} - [y - (z_1 + z_2)]$$

is an exponential observer with $z_1(0)$, $z_2(0)$ arbitrarily given for the system of the example.

Thau's Observers @ 1975

Consider the class of autonomous nonlinear systems

$$\Sigma_{\mathcal{G}}: \begin{cases} \dot{x} = Ax + \phi(x, y, \dot{y}) \\ y = Cx \end{cases} \qquad \begin{matrix} x : x(t) \in R^n \\ y : y(t) \in R^p \end{matrix} \qquad (9)$$

where A and C are known and $\phi(.)$ is Lipschitz in its arguments and verifies

$$\phi(x, y, \dot{y}) = \phi_1(y) + \nabla\phi_2(y)\dot{y} + \phi_3(x) \qquad (10)$$

where $\phi_1(.), \phi_3(.) \in C^1, \phi_2(.) \in C^2$ and such that

$$C\nabla\phi_2(y)\dot{y} = 0.$$

Exponential Observers @ 1975

Theorem[3] : For the nonlinear autonomous systems (9), verifying (10) if

a) the pair (A,C) is observable

b) there exist $P > 0,\ Q > 0$, and a gain vector K

such that

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

and

$$\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} > \|\nabla\phi_3\|_\infty \qquad (11)$$

then there exist an exponential observer for system (9).

Elements of Proof (I)

Define the new variable $w = x - \phi_2(\mathrm{y})$, and note that

$$\dot{\phi}_2 = \nabla\phi_2(y)\dot{y}$$

The derivative, relative to time, of this new signal verifies

$$\dot{w} = Ax + \phi_1(y) + \nabla\phi_2(y)\dot{y} + \phi_3(x) - \nabla\phi_2(y)\dot{y}$$

That can be simplified to

$$\dot{w} = Aw + \phi_1(y) + A\phi_2(y) + \phi_3(w + \phi_2)$$

Considering that $y_1 = Cw = y - C\phi_2(y)$, the observer

$$\dot{\hat{w}} = A\hat{w} + \phi_1(y) + A\phi_2(y) + \phi_3(\hat{w} + \phi_2) - K(y_1 - C\hat{w})$$

is proposed.

Elements of Proof (II)

Defining $e = w - \hat{w}$, the error dynamics is given by

$$\dot{e} = (A - KC)e + \phi_3(w + \phi_2) - \phi_3(\hat{w} + \phi_2)$$

For the Lyapunov candidate function $V(t) = e^T P e$,

where $P > 0$ is a symmetric constant matrix (2^{nd} method)

$$\dot{V}(t) = \dot{e}^T P e + e^T P \dot{e}$$

$$= -2e^T Q e +$$

$$(\phi_3(w + \phi_2) - \phi_3(\hat{w} + \phi_2))^T P e + e^T P(\phi_3(w + \phi_2) - \phi_3(\hat{w} + \phi_2))$$

Elements of Proof (III)

$$
\begin{aligned}
\dot{V} &\leq -2\lambda_{\min}(Q) + \left(\int_0^1 \nabla \phi_3 (w_s + \phi_2(y)) ds e \right)^T Pe \\
&\quad + e^T P \left(\int_0^1 \nabla \phi_3 (w_s + \phi_2(y)) ds e \right) \\
&\leq \left[-2\lambda_{\min}(Q) + 2\|P\| \|\nabla \phi_3\|_\infty \right] \|e\|^2
\end{aligned}
$$

From this relation (11) is immediate.

Example, from [3]

Example 3. We consider a simple pendulum with viscous damping and without driving torque.

$$\ddot{x} + a_2\dot{x} + a_3 \sin x = 0, \qquad y = x, \tag{28}$$

where a_2 , a_3 are constants.

Let us rewrite Eq. (28) as a vector differential equation

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ 0 & -a_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ -a_3 \sin x_1 \end{bmatrix},$$

$$y = [1 \;\; 0]\, \mathbf{x}.$$

Now the linear part of this system is observable and we denote the nonlinear part by

$$\phi(\mathbf{x}) = \begin{bmatrix} 0 \\ -a_3 \sin x_1 \end{bmatrix} \quad \text{so} \quad \nabla\phi(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ -a_3 \cdot \cos x_1 & 0 \end{bmatrix}.$$

If $a_2 = \frac{3}{4}$, $a_3 = \frac{1}{8}$ then the following matrices

$$B = \begin{bmatrix} \frac{3}{4} \\ -\frac{1}{4} \end{bmatrix} \qquad Q = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{and} \qquad P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{29}$$

satisfy Condition (b) of Theorem 3. So the observer (23) with B given by (29) and $\phi_1 = 0$, $\phi_2 = 0$ is an exponential observer and the ϵ of (27) is $\frac{1}{8}(5 - (5)^{1/2}) > 0$.

Lipschitz Observers @ 1998

Consider the class of non-autonomous nonlinear systems

$$\Sigma_{\mathcal{G}} : \begin{cases} \dot{x}(t) = Ax(t) + \phi(x(t), u(t)) \\ y(t) = Cx(t) \end{cases} \qquad \begin{matrix} x(t) \in R^n \\ y(t) \in R^p \\ u(t) \in R^m \end{matrix} \qquad (12)$$

A, C, and $f(.)$ are known; (A, C) is observable, u is a deterministic input, and $\phi(.)$ is a Lipschitz time-invariant function, i.e.

$$\left\| \phi(x(t), u(t)) - \phi(\hat{x}(t), u(t)) \right\| \leq \gamma \left\| x(t) - \hat{x}(t) \right\| \qquad (13)$$

Proposed observer (motivated by Luenberger's and Kalman's work)

$$\dot{\hat{x}}(t) = A\hat{x}(t) + \phi(\hat{x}(t), u(t)) + K(y(t) - C\hat{x}(t)) \qquad (14)$$

Lipschitz Observers @ 1998

The estimation error dynamics is

$$\dot{\tilde{x}}(t) = (A - KC)\tilde{x}(t) + [\phi(x(t), u(t)) - \phi(\hat{x}(t), u(t))]$$

and the following major result holds:

Theorem [5] : For the nonlinear non - autonomous systems (12), verifying (13), if the observer given by (14) satisfies

a) the pair *(A,C)* is observable

b) the gain *K* can be chosen such that as to ensure

$$\min_{\omega \in R^+} \sigma_{\min}(A - KC - Ij\omega) > \gamma \qquad (15)$$

then it is asymptotically stable.

Elements of Proof (I)

The proof is done in three parts (see [5] for details):

i)

If $\min\limits_{\omega \in R^+} \sigma_{\min}\left(A - KC - j\omega I\right) > \gamma$, then there exits $\varepsilon > 0$ such that the matrix

$$H = \begin{bmatrix} (A\text{-}KC) & \gamma^2 I \\ -I - \varepsilon I & -(A\text{-}KC)^T \end{bmatrix}$$

has no imaginary eigenvalues.

ii)

If $(A - KC)$ is stable then there exists a $P > 0$ such that there exits a solution to the equation

$$(A - KC)^T P + P(A - KC) + \gamma^2 PP + I + \varepsilon I = 0$$

Elements of Proof (II)

cont...

iii) Defining $\tilde{x}(t) = x(t) - \hat{x}(t)$, the error dynamics is given by

$$\dot{\tilde{x}}(t) = (A - KC)\tilde{x}(t) + [\phi(x(t), u(t)) - \phi(\hat{x}(t), u(t))]$$

For the Lyapunov candidate function $V(t) = \tilde{x}(t)^T P \tilde{x}(t)$,

where $P > 0$ is a symmetric constant matrix (2^{nd} method)

$$\dot{V}(t) = \frac{d}{dt}\tilde{x}(t)^T P \tilde{x}(t) = \dot{\tilde{x}}(t)^T P \tilde{x}(t) + \tilde{x}(t)^T P \dot{\tilde{x}}(t)$$

$$= (A - KC)^T P \tilde{x} + \tilde{x}^T P (A - KC) +$$

$$2\tilde{x}^T P[\phi(x(t), u(t)) - \phi(\hat{x}(t), u(t))]$$

Using the properties introduced before, results:

$$\dot{V}(t) \le \tilde{x}^T \left[(A - KC)^T P + P(A - KC) + \gamma^2 PP + I \right] \tilde{x}$$

References

[1] P. Kokotovic , *The Joy of Feedback: Nonlinear and Adaptive,* Macmillan Press, 1991 Bode Prize Lecture, 1992.

[2] F. Thau, *”Observing the state of nonlinear dynamical systems,”* International Journal of Control, pp. 471-479, 1973.

[3] S. Kou, D. Elliot, and T. Tarn, ”Exponential observers for nonlinear dynamic systems,” Journal of Information and Control pp. 204-216, 1975.

[4] Wen Yu, *”Nonlinear Observer,”* class notes, CINVESTAV-IPN.

[5] R. Rajamani, ”Observers for Lipschitz Nonlinear Systems,” Transactions on Automatic Control, pp. 397-400, 1998.