Nonlinear optimization-based state estimation: robustness analysis by Q functions

James B. Rawlings and Douglas A. Allan

Department of Chemical Engineering

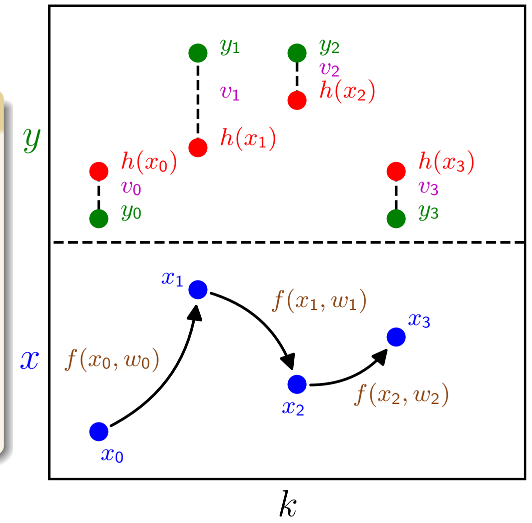
Monterey Workshop on Computational Control
Monterey, CA
October 7–8, 2019

Optimization-based state estimation—Introduction

System model

$$x^+ = f(x, w)$$
$$y = h(x) + v$$

State	$x \in \mathbb{R}^n$
Output	$y \in \mathbb{R}^p$
Process noise	$w \in \mathbb{R}^g$
Output noise	$v \in \mathbb{R}^p$



Full information estimation—optimal control problem

- Solve a nonlinear program with objective function $V_k(\cdot)$ that, e.g., serves as a surrogate for the likelihood function

$$\min_{\chi(0),\boldsymbol{\omega}} V_k(\chi(0), \boldsymbol{\omega}; \overline{x}_0, \mathbf{y}) := |\chi(0) - \overline{x}_0|^2_{P_0^{-1}} + \sum_{j=0}^{k-1} |\omega(j)|^2_{Q^{-1}} + |\nu(j)|^2_{R^{-1}}$$

$$\text{subject to:} \qquad \chi^+ = f(\chi, \omega) \qquad y = h(\chi) + \nu$$

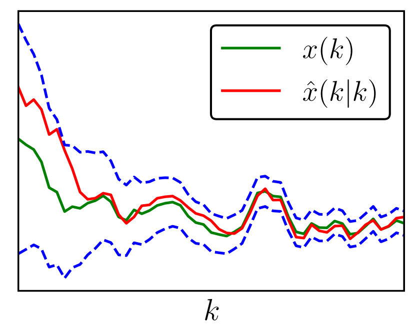
- For each time $k \in \mathbb{I}_{\geq 0}$, a sequence of state estimates $(\hat{x}(0|k), \hat{x}(1|k), \ldots, \hat{x}(k|k))$, state disturbance estimates $(\hat{w}(0|k), \ldots, \hat{w}(k-1|k))$, and measurement disturbance estimates $(\hat{v}(0|k), \ldots, \hat{v}(k-1|k))$ are generated, with optimal cost $V_k^0(\overline{x}_0, \mathbf{y})$

Robust stability of estimation—the steady-state Kalman filter—what do we know

- Linear system.

$$x^+ = Ax + Gw$$
$$y = Cx + v$$

- If (A, C) is *detectable*, (A, G) is stabilizable, $P_0, Q, R > 0$, then steady-state KF error satisfies



$$\boxed{|x_k - \hat{x}_k| \leq c\,|x_0 - \overline{x}_0|\,\lambda^k + \frac{c\,|G|}{1-\lambda}\,\|\mathbf{w}\|_{0:k-1} + \frac{c\,|L|}{1-\lambda}\,\|\mathbf{v}\|_{0:k-1}}$$

using the sup norm over the sequence, $\|\mathbf{w}\|_{0:k-1} := \max_{j \in 0:k-1} |w(j)|$

- The goal of this talk is to extend this robust exponential stability result to the nonlinear case

Assumption: Regularity

Assumption 1 (Continuity)

The functions $f(\cdot)$ and $h(\cdot)$ are continuous.

Assumption 2 (Positive definite costs)

The matrices Q, R, and P_0 are positive definite.

- Under these assumptions, the FIE problem has a solution for all $k \in \mathbb{I}_{\geq 0}$.

Assumption: Detectability

Definition 3 (Exponential incremental input/output-to-state stability)

A system is exponentially incrementally input/output-to-state stable (exp i-IOSS) if there exist $\lambda \in (0,1)$ and $c_x, c_w, c_y > 0$ such that

$$\begin{aligned}|x_1(k) - x_2(k)| \leq c_x |x_1(0) - x_2(0)| \lambda^k + c_w \|\mathbf{w}_1 - \mathbf{w}_2\|_{0:k-1} \\ + c_y \|\mathbf{y}_1 - \mathbf{y}_2\|_{0:k-1}\end{aligned}$$

for every $x_1(0), x_2(0) \in \mathbb{X}$, $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}$, and $k \in \mathbb{I}_{\geq 0}$.

Definition 4 (Exponential i-IOSS Lyapunov function)

A function $\Lambda : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{\geq 0}$ is an exp i-IOSS Lyapunov function if there exist $\sigma > 0$ and $c_1, c_2, c_3, c_w, c_y > 0$ such that

$$c_1 |x_1 - x_2|^\sigma \leq \Lambda(x_1, x_2) \leq c_2 |x_1 - x_2|^\sigma$$
$$\Lambda(x_1^+, x_2^+) \leq \Lambda(x_1, x_2) - c_3 |x_1 - x_2|^\sigma + c_w |w_1 - w_2|^\sigma + c_y |y_1 - y_2|^\sigma$$

Assumption: Exponential Detectability

Theorem 5

A system is exp i-IOSS if and only if it admits an exp i-IOSS Lyapunov function.

Assumption 6 (Detectability)

The system is exp i-IOSS, and thus admits an exp i-IOSS Lyapunov function of the form

$$c_1 |x_1 - x_2|^2 \leq \Lambda(x_1, x_2) \leq c_2 |x_1 - x_2|^2$$
$$\Lambda(x_1^+, x_2^+) \leq \Lambda(x_1, x_2) - c_3 |x_1 - x_2|^2 + |w_1 - w_2|^2_{Q^{-1}} + |y_1 - y_2|^2_{R^{-1}}$$

in which Q and R come from the stage cost.

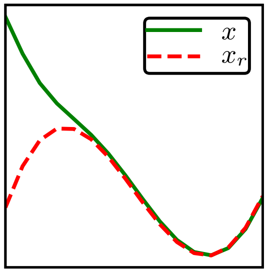
Assumption: Incremental Stabilizability

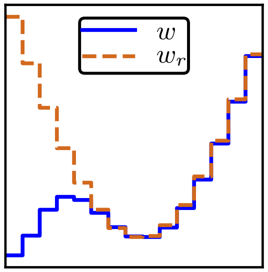
Assumption 7 (Stabilizability)

The system $x^+ = f(x, w)$ is incrementally exponentially stabilizable if there exists $\bar{c} > 0$ such that for every initial state $x \in \mathbb{X}$, every reference state $x_r \in \mathbb{X}$, and reference input sequence $\mathbf{w}_r$, there exists an input sequence $\mathbf{w}$ such that

$$\sum_{k=0}^{\infty} |w(k) - w_r(k)|^2_{Q^{-1}} + |y(k) - y_r(k)|^2_{R^{-1}} \leq \bar{c} |x - x_r|^2$$

in which $x^+ = f(x, w)$, $y = h(x)$, $x_r^+ = f(x_r, w_r)$, and $y_r = h(x_r)$.





Comparison of assumptions to past work

- The sort of system regularity assumptions are standard in optimization-based state estimation literature
- i-IOSS is a common nonlinear detectability assumption
- The introduction of the i-IOSS Lyapunov function as an analysis tool for FIE is a novel contribution
- Stabilizability has been largely absent from prior work on nonlinear FIE and MHE, but the use of an additive state disturbance, $x^+ = f(x) + w$, can be viewed as a tacit, unnecessarily strong assumption of controllability

Nominal Stability

Definition 8 (Exponentially Stable Estimator)

A state estimator is exponentially stable if there exists $\lambda \in (0,1)$ and $C > 0$ such that its estimates $\hat{x}(k)$ satisfy

$$|\hat{x}(k) - x(k)| \leq C\,|\overline{x}_0 - x(0)|\,\lambda^k$$

for all k, in which $\overline{x}_0$ is the prior information on the initial state.

- To present the new analysis, we first consider the nominal stability of FIE, i.e., when $w(k) = v(k) = 0$ for all $k \in \mathbb{I}_{\geq 0}$, but $\overline{x}_0 \neq x(0)$
- Analysis of this type first appeared in Allan and Rawlings (2019)

Infinite-horizon problem

- The choice $\chi(0) = x(0)$, $\omega(j) = 0$, and $\nu(j) = 0$ is feasible for the FIE problem at time k, and as a result $V_k^0 \leq |\overline{x}_0 - x(0)|^2_{P_0^{-1}}$ for all k (the dependence of $V_k^0(\cdot)$ on $\overline{x}_0$ and $\mathbf{y}$ has been suppressed for brevity)
- Because the sequence $(V_1^0, V_2^0, \ldots)$ is nondecreasing and bounded above, it converges to some $V_\infty^0 \leq |\overline{x}_0 - x(0)|^2_{P_0^{-1}}$
- It can be shown (Keerthi and Gilbert, 1985) that there exists some sequences $\widehat{\mathbf{x}}(\infty)$, $\widehat{\mathbf{w}}(\infty)$, and $\widehat{\mathbf{v}}(\infty)$ such that

$$V_\infty(\widehat{\mathbf{x}}(\infty), \widehat{\mathbf{w}}(\infty), \widehat{\mathbf{v}}(\infty)) = V_\infty^0$$

Turning the problem on its head

- In standard regulation, we have a sequence of optimal costs that are *nonincreasing* and bounded *below*
- In FIE, we instead have a *nondecreasing* sequence bounded *above*
- If we define

$$Z(k) := V_\infty^0 - V_k^0$$

  then we have a sequence nonincreasing and convergent to zero
- By the principle of optimality,

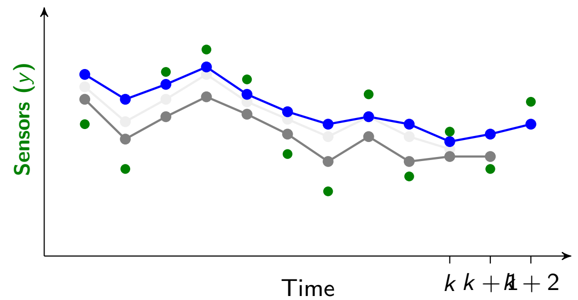
$$V_k^0 \leq V_{k+1}^0 - |\hat{w}(k|k+1)|^2_{Q^{-1}} - |\hat{v}(k|k+1)|^2_{R^{-1}}$$

  so we have

$$Z(k+1) - Z(k) = V_k^0 - V_{k+1}^0 \leq -|\hat{w}(k|k+1)|^2_{Q^{-1}} - |\hat{v}(k|k+1)|^2_{R^{-1}}$$

  as a descent condition
- As a result, we know that $\hat{w}(k|k+1)$ and $\hat{v}(k|k+1)$ converge to zero

Inadequacy of descent condition



- There are two problems using this descent condition
  1. No rate of convergence is given for $\hat{w}(k|k+1)$ and $\hat{v}(k|k+1)$
  2. $\hat{x}(k+1) \neq f(\hat{x}(k), \hat{w}(k|k+1))$
- For every new measurement, the entire state trajectory must be reconstructed
- In order to apply detectability condition, a trajectory must satisfy the system evolution equation $x^+ = f(x, w)$

Second time index

- Alternative: compute a cost decrease condition *within* a trajectory $\widehat{\mathbf{x}}(k)$
- Introduce partial sum of trajectory

$$V^0(j|k) := |\hat{x}(0|k) - \overline{x}_0|^2_{P_0^{-1}} + \sum_{i=0}^{j-1} |\hat{w}(i|k)|^2_{Q^{-1}} + |\hat{v}(i|k)|^2_{R^{-1}}$$

- Similarly, add the second time index j to $Z(\cdot)$

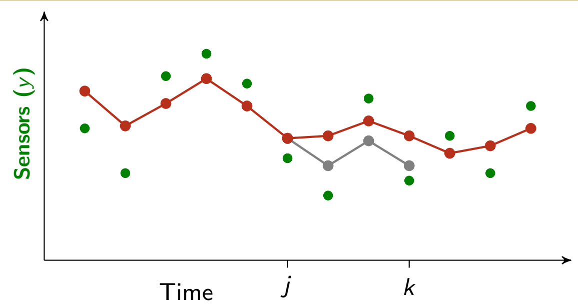
$$Z(j|k) := V^0_\infty - V^0(j|k)$$

- Obtaining a descent condition in j is easy

$$Z(j+1|k) - Z(j|k) = V^0(j|k) - V^0(j+1|k) = -|\hat{w}(j|k)|^2_{Q^{-1}} - |\hat{v}(j|k)|^2_{R^{-1}}$$

(Note: an equality, not even an inequality)

Tail reoptimization



- Obtain upper bound for V^0_∞ by reoptimizing trajectory from $\hat{x}(j|k)$

$$V^0_\infty \leq V^0(j|k) + \min_{\omega,\nu} \sum_{i=j}^{\infty} |\omega(i)|^2_{Q^{-1}} + |\nu(i)|^2_{R^{-1}}$$
$$\text{subject to} \quad \chi(i+1) = f(\chi(i), \omega(i))$$
$$y(i) = h(\chi(i)) + \nu(i)$$
$$\chi(j) = \hat{x}(j|k)$$

Optimal control problem

- Let's examine the optimization problem

$$\min_{\omega,\nu} \sum_{i=j}^{\infty} |\omega(i)|^2_{Q^{-1}} + |\nu(i)|^2_{R^{-1}}$$
$$\text{subject to} \quad \chi(i+1) = f(\chi(i), \omega(i))$$
$$\chi(j) = \hat{x}(j|k) \qquad y(i) = h(\chi(i)) + \nu(i)$$

- Because $\chi(j)$ is not a degree of freedom, it is an infinite-horizon tracking problem with initial state $\hat{x}(j|k)$
- The stabilizability assumption gives an upper bound

$$\min_{\omega,\nu} \sum_{i=j}^{\infty} |\omega(i)|^2_{Q^{-1}} + |\nu(i)|^2_{R^{-1}} \leq \overline{c}\, |\hat{x}(j|k) - x(j)|^2$$

Upper and lower bounds

- We thus have that

$$Z(j|k) = V^0_\infty - V^0(j|k) \leq \bar{c}\,|\hat{x}(j|k) - x(j)|^2$$

- Furthermore, because $V^0(j|k) \leq V^0(k|k) \leq V^0_\infty$, we have that

$$Z(j|k) \geq 0$$

- $Z(\cdot)$ has a semidefinite lower bound and semidefinite cost decrease

$$Z(j+1|k) - Z(j|k) = V^0(j|k) - V^0(j+1|k) = -\,|\hat{w}(j|k)|^2_{Q^{-1}} - |\hat{v}(j|k)|^2_{R^{-1}}$$

Applying *detectability assumption*

- The i-IOSS Lyapunov function allows us to turn these semidefinite bounds to fully definite bounds (see Grimm et al. (2005) for a similar idea in regulation). Let

$$Q(j|k) := \Lambda(\hat{x}(j|k), x(j)) + Z(j|k)$$

- We immediately have

$$c_1\,|\hat{x}(j|k) - x(j)|^2 \leq Q(j|k) \leq \bar{c}_2\,|\hat{x}(j|k) - x(j)|^2$$

in which $\bar{c}_2 := c_2 + \bar{c}$

Strict descent

- Finally, the supply rate in the dissipation inequality

$$\begin{aligned}\Lambda(\hat{x}(j+1|k), x(j+1)) \leq{}& \Lambda(\hat{x}(j|k), x(j)) - c_3\,|\hat{x}(j|k) - x(j)|^2 \\ &+ |\hat{w}(j|k)|^2_{Q^{-1}} + |\hat{v}(j|k)|^2_{R^{-1}}\end{aligned}$$

cancels with the descent condition for $Z(\cdot)$

$$Z(j+1|k) \leq Z(j|k) - |\hat{w}(j|k)|^2_{Q^{-1}} - |\hat{v}(j|k)|^2_{R^{-1}}$$

to obtain

$$Q(j+1|k) \leq Q(j|k) - c_3\,|\hat{x}(j|k) - x(j)|^2$$

- Now $Q(\cdot)$ looks like a full Lyapunov function

Standard Lyapunov argument

- By combining upper bound on $Q(\cdot)$ and descent condition, we obtain

$$Q(j+1|k) \leq Q(j|k) - (c_3/\bar{c}_2)Q(j|k)$$

- Let $\sigma := (1 - c_3/\bar{c}_2) \in (0, 1)$.

$$Q(j+1|k) \leq \sigma(Q(j|k))$$

- Iterate to obtain

$$Q(k|k) \leq \sigma^k Q(0|k)$$

- Thus, by applying the lower and upper bounds, we have that

$$|\hat{x}(k|k) - x(k)| \leq \sqrt{\bar{c}_2/c_1}\,|\hat{x}(0|k) - x(0)|\,\sigma^k$$

Uniform upper bound for $|\hat{x}(0|k) - x(0)|$

- Unfortunately $\hat{x}(0|k)$ is recalculated at every time k
- So need an upper bound for $|\hat{x}(0|k) - x(0)|$
- Let $\overline{\lambda}_0$ and$\underline{\lambda}_0$ be the largest and smallest eigenvalues of P_0, respectively. For any vector a

$$(1/\overline{\lambda}_0)\,|a|^2 \leq |a|^2_{P_0^{-1}} \leq (1/\underline{\lambda}_0)\,|a|^2$$

- Because $\chi(0) = x(0)$, $\boldsymbol{\omega} = \boldsymbol{\nu} = 0$ is feasible

$$\frac{1}{\overline{\lambda}_0}\,|\hat{x}(0|k) - \overline{x}_0|^2 \leq |\hat{x}(0|k) - \overline{x}_0|^2_{P_0^{-1}} \leq V_k^0 \leq V_\infty^0 \leq \frac{1}{\underline{\lambda}_0}\,|\overline{x}_0 - x(0)|^2$$

- Apply the triangle inequality

$$|\hat{x}(0|k) - x(0)| \leq |\hat{x}(0|k) - \overline{x}_0| + |\overline{x}_0 - x(0)|$$

- To obtain

$$\boxed{|\hat{x}(k|k) - x(k)| \leq \sqrt{\overline{c}_2/c_1}\left(1 + \sqrt{\kappa(P_0)}\right)|\overline{x}_0 - x(0)|\,\sigma^k}$$

and FIE is exponentially stable.

Q-function summary

Definition 9

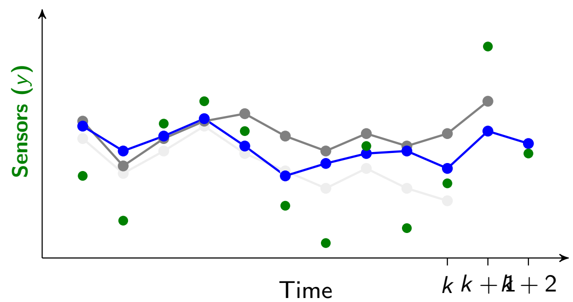
A function $Q(j|k; \overline{x}_0, \mathbf{y})$ is an exponential *Q-function* if there exist $C_0, C_1, C_2, C_3 > 0$ such that

$$Q(0|k) \leq C_0\,|\overline{x}_0 - x(0)|^2 \qquad k \in \mathbb{I}_{\geq 0}$$
$$C_1\,|\hat{x}(j|k) - x(j)|^2 \leq Q(j|k) \leq C_2\,|\hat{x}(j|k) - x(j)|^2 \quad j \leq k \in \mathbb{I}_{\geq 0}$$
$$Q(j+1|k) \leq Q(j|k) - C_3\,|\hat{x}(j|k) - x(j)|^2 \quad j \leq k - 1 \in \mathbb{I}_{\geq 0}$$

Theorem 10

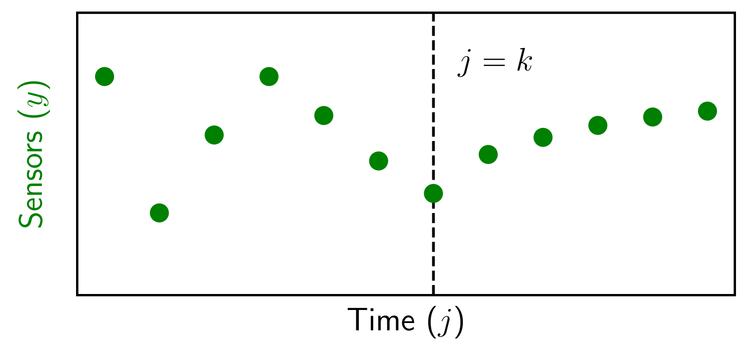
If a state estimation scheme admits an exponential Q-function, then it is exponentially stable.

From stability to robustness



- The main difference between regulation and estimation is that the entire state trajectory is reestimated at every time
- This difference matters even more when disturbances are involved—future disturbances can throw off smoothed estimates of past states

Infinite-horizon cost—what now?



- With persistent disturbances, there is no bounded infinite horizon cost
- Instead, base a sequence of infinite horizon costs $V_\infty^0(k)$ on a sequence of outputs $\tilde{\mathbf{y}}(k)$ in which disturbances end at time k

Upper bound with disturbances

- In addition to accommodating the estimation error $\hat{x}(j|k) - x(j)$, we must also accommodate upcoming disturbances
- Leads to upper bounds

$$Q(0|k) \leq C_{0,x} \left|x(0) - \overline{x}_0\right|^2 + C_{0,d} \sum_{i=0}^{k-1} |w(i)|^2_{Q^{-1}} + |v(i)|^2_{R^{-1}}$$

for initial cost and

$$Q(j|k) \leq C_{2,x} \left|\hat{x}(j|k) - x(j)\right|^2 + C_{2,d} \sum_{i=j}^{k-1} |w(i)|^2_{Q^{-1}} + |v(i)|^2_{R^{-1}}$$

for subsequent costs

Cost decrease becomes dissipation

- Because we now have disturbances, we no longer have a simple cost decrease, but a dissipation inequality with a supply rate
$$Q(j+1|k) \leq Q(j|k) - C_{3,x} \left|\hat{x}(j|k) - x(j)\right|^2 + C_{3,w} |w(j)|^2 + C_{3,v} |v(j)|^2$$

Definition 11

A function $Q(j|k; \overline{x}_0, \mathbf{y})$ is a robust exponential *Q-function* if there exist $C_0, C_1, C_2, C_3 > 0$ such that

$$Q(0|k) \leq C_{0,x} \left|\overline{x}_0 - x(0)\right|^2 + C_{0,d} \sum_{i=0}^{k-1} |w(i)|^2_{Q^{-1}} + |v(i)|^2_{R^{-1}}$$

$$Q(j|k) \geq C_1 \left|\hat{x}(j|k) - x(j)\right|^2$$

$$Q(j|k) \leq C_{2,x} \left|\hat{x}(j|k) - x(j)\right|^2 + C_{2,d} \sum_{i=j}^{k-1} |w(i)|^2_{Q^{-1}} + |v(i)|^2_{R^{-1}}$$

$$Q(j+1|k) \leq Q(j|k) - C_{3,x} \left|\hat{x}(j|k) - x(j)\right|^2 + C_{3,w} |w(j)|^2 + C_{3,v} |v(j)|^2$$

Robustness result

Theorem 12

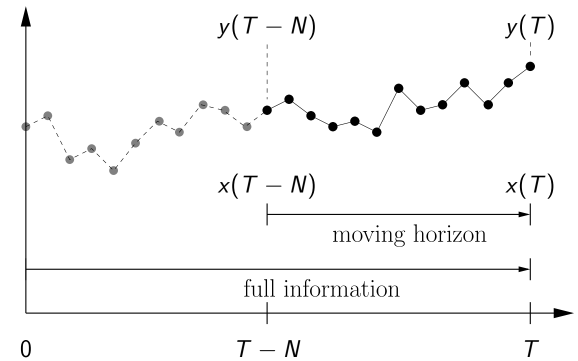
If an estimator admits a robust exponential Q function, then it is robustly exponentially stable, i.e., there exist $C_x, C_w, C_v > 0$ and $\lambda \in (0,1)$ such that

$$\left|\hat{x}(k|k) - x(k)\right| \leq C_x \left|\overline{x}_0 - x(0)\right| \lambda^k + C_w \|\mathbf{w}\|_{0:k-1} + C_v \|\mathbf{v}\|_{0:k-1}$$

for all $k \geq 0$.

- Note that this is the same(!) result as the steady-state Kalman filter for $f(x,w) = Ax + Gw$ with (A,C) detectable and (A,G) stabilizable.
- This result does *not* guarantee anything about the asymptotic behavior of the smoothed estimates $\hat{x}(j|k)$ for all $j \leq k$
- But it does cover stability of the fixed-lag smoother $\hat{x}(k-p|k)$ for any fixed $p \leq k$.

Application to MHE



- In MHE, only the N most recent measurements are used

$$\min_{\chi(k-N),\boldsymbol{\omega},\boldsymbol{\nu}} V_k := \left|\chi(k-N) - \overline{x}(k-N)\right|^2_{P^{-1}_{k-N}} + \sum_{j=k-N}^{k-1} |\omega(j)|^2_{Q^{-1}} + |\nu(j)|^2_{R^{-1}}$$

subject to $\quad \chi^+ = f(\chi, \omega) \qquad y = h(\chi) + \nu$

Previous results for MHE

- MHE was shown to be robustly stable for observable systems in Rao et al. (2003)
  - Problem: bound on $|\hat{x}(k|k) - x(k)|$ gets worse with increasing horizon length N
- For locally exponentially detectable systems, MHE was shown to be robustly stable in Müller (2017)
  - Requires several difficult-to-interpret assumptions
  - Because bound on $|\hat{x}(k|k) - x(k)|$ gets worse with increasing horizon length N, a delicate balancing act is necessary when choosing estimator design parameters
- Special case of MHE was shown to be robustly stable for exponentially detectable systems in Knüfer and Müller (2018)
  - This result depends on both exponentially discounting past measurements and using an ℓ_1 cost function
  - Also requires additive w

Results for MHE

Theorem 13

If MHE is performed on a system satisfying our assumptions with a filtering prior and constant weighting $P_{k-N} = P$, there exists a horizon N^* such that, if $N \geq N^*$, MHE is robustly stable.

- In short—for exponentially detectable systems, MHE works so long as the horizon is long enough
- Critically, the bound on $|\hat{x}(k|k) - x(k)|$ gets *better* with increasing horizon length N
- Areas of further research:
  - Smart ways of updating P_{k-N} to shorten the horizon and reduce online computation
  - Extension to asymptotically detectable MHE and more general stage cost

References

D. A. Allan and J. B. Rawlings. A Lyapunov-like function for full information estimation. In *American Control Conference*, pages 4497–4502, Philadelphia, PA, July 10–12, 2019.

G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel. Model predictive control: For want of a local control Lyapunov function, all is not lost. *IEEE Trans. Auto. Cont.*, 50(5):546–558, 2005.

S. S. Keerthi and E. G. Gilbert. An existence theorem for discrete-time infinite-horizon optimal control problems. *IEEE Trans. Auto. Cont.*, 30(9): 907–909, September 1985.

S. Knüfer and M. A. Müller. Robust Global Exponential Stability for Moving Horizon Estimation. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 3477–3482, Dec 2018.

M. A. Müller. Nonlinear moving horizon estimation in the presence of bounded disturbances. *Automatica*, 79:306–314, 2017.

C. V. Rao, J. B. Rawlings, and D. Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations. *IEEE Trans. Auto. Cont.*, 48(2):246–258, February 2003.