Guaranteed State Estimation for Linear Time-Varying Systems *

Jeff S. Shamma Edward Tu
Department of Aerospace Engineering and Engineering Mechanics
The University of Texas at Austin
Austin, TX 78712

Abstract

A guaranteed state estimator produces a set of possible states based on output measurements and models of exogenous signals. In this paper, we consider the guaranteed state estimation problem for linear time-varying systems with *a priori* magnitude bounds on exogenous signals. We provide a recursive algorithm to propagate the set of possible states based on output measurements. We then show that the center of the sets provides an optimal estimate in an ℓ^∞-induced norm sense.

1 Introduction

Stochastic state estimation provides optimal state estimates based on probabilistic models of exogenous signals. An alternative is to model exogenous signals as deterministic unknown but bounded quantities. The problem is then to construct a *set* of possible state values based on measured outputs. Such an approach has received considerable attention in the controls literature. References [1, 5] present an overview of work in this area, and reference [4] contains a collection of related conference papers.

Related to the deterministic setting is the induced-norm optimal state estimation. This framework provides optimal state estimates which minimize the induced-norm from exogenous signals to estimation errors. Reference [6] considers the case where exogenous signals and estimation errors are measured using the ℓ^2-norm, or signal energy, which leads to an $\mathcal{H}^\infty$ optimal estimation problem. Reference [7] measures exogenous signals and estimation errors by the ℓ^∞ norm, or signal magnitude, which leads to an ℓ^1 optimal estimation problem.

In this paper, we consider guaranteed state estimation for linear time-varying systems. Under an assumed *a priori* bound on exogenous signals, we present a recursive construction of the set of possible state values. We then relate the center of these sets to the ℓ^1 optimal estimation problem considered in [7]. In particular, we show that the centers are also optimal in an ℓ^∞ induced norm sense.

*Supported by NSF Grant ECS-9258005 and Ford Motor Co.

The remainder of this paper is organized as follows. Section 2 contains preliminary definitions and notation. Section 3 presents an algorithm which propagates the set-valued estimates based on output measurements. The main results are in Section 4 which derives the ℓ^∞ induced-norm optimality of the centers of these sets. Finally, Section 5 has concluding remarks.

2 Notation

Let $\mathcal{Z}^+$ denote the set of non-negative integers. For $x \in \mathcal{R}^n$, let x_i denote the i^{th} component of x and define $|x| = \max_i |x_i|$. Let ℓ^∞ denote the set of bounded one-sided sequences in $\mathcal{R}^n$. For $f = \{f(0), f(1), f(2), \ldots\} \in \ell^\infty$, define

$$\|f\| = \sup_{k \in \mathcal{Z}^+} |f(k)| .$$

The dimension n is suppressed in ℓ^∞ for notational convenience. The unit balls in $\mathcal{R}^n$ and ℓ^∞ are denoted $B_{\mathcal{R}^n}$ and B_{ℓ^∞}, respectively. The truncation operator $P_N : \ell^\infty \to \ell^\infty$ is defined by

$$P : \{f(0), f(1), f(2), \ldots\} \mapsto \{f(0), f(1), \ldots, f(N), 0, \ldots\} .$$

For $M \in \mathcal{R}^{z \times n}$ and $m \in \mathcal{R}^z$, let $Set(M, m)$ denote the subset of $\mathcal{R}^n$ associated with (M, m) defined by the constraints

$$Set(M, m) = \{x : Mx \leq m\} .$$

For $M_1 \in \mathcal{R}^{z \times n}$, $M_2 \in \mathcal{R}^z$, and $m \in \mathcal{R}^z$ consider the subset, S, of $\mathcal{R}^n$ defined by

$$S = \{x : M_1 x + M_2 w \leq m \text{ for some } w \in \mathcal{R}\} .$$

Define

$$\begin{aligned} &Rack\,[(\, M_1 \quad M_2\,)\,, m] \\ &\quad = \left\{ (\tilde{M}, \tilde{m}) \in \mathcal{R}^{z \times n} \times \mathcal{R}^z : S = Set(\tilde{M}, \tilde{m}) \right\} , \end{aligned}$$

i.e., $Rack\,[(\, M_1 \quad M_2\,)\,, m]$ is the set of matrix pairs which give a direct characterization of S. The construction of an element of $Rack\,[(\, M_1 \quad M_2\,)\,, m]$ may be achieved through the Fourier-Motzkin algorithm which is described in [3] and reviewed in the Appendix.

3 Set-Valued Estimation

We consider the time-varying discrete-time linear system,

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)d(k), \quad x(0) = x_o, \\ y(k) &= C(k)x(k) + n(k), \end{aligned} \qquad (1)$$

where x is the state-vector, y is the measured output, d is a process disturbance, and n is a measurement noise. Define $w = \begin{pmatrix} d \\ n \end{pmatrix}$. In input/output form, the system (1) takes the form,

$$y = T_{yw}w + T_{yx_o}x_o,$$

where T_{yw} denotes the mapping from w to y with the initial condition $x_o = 0$, and T_{yx_o} denotes the mapping from x_o to y with the input $w = 0$. Similarly define T_{xw} and T_{xx_o}.

Assumption 3.1 *We make the following assumptions on (1) throughout.*

a) $T_{xw} : \ell^\infty \to \ell^\infty$ *and* $T_{xx_o} : \mathcal{R}^n \to \ell^\infty$ *are bounded linear operators.*

b) $A(k)$ *is invertible for all* $k \in \mathcal{Z}^+$.

c) $w \in B_{\ell^\infty}$ *and* $x(0) \in B_{\mathcal{R}^n}$.

d) d, n, *and* y *are scalar-valued signals.*

Under appropriate observability conditions, Assumptions 3.1a–b are made with minimal loss of generality. In the time-invariant case, they can always be achieved through the use of a conventional observer as a kind of preliminary feedback. Assumption 3.1c represents the *a priori* model on exogenous signals initial conditions. Assumption 3.1d is for notational convenience. The methods considered here easily extend to multivariable systems.

We are interested in constructing an estimate of the state vector based on output measurements. Towards this end, define the set-valued map $W : \ell^\infty \times \mathcal{Z}^+ \rightsquigarrow \mathcal{R}^n$ as

$$\begin{aligned} W(y,N) = \{ & (w, x_o) \in B_{\ell^\infty} \times B_{\mathcal{R}^n} : \\ & P_N y = P_N(T_{yw}w + T_{yx_o}x_o) \}. \end{aligned}$$

In other words, $W(y,N)$ denotes the set of admissible exogenous signals and initial conditions consistent with measured data up to time n. Similarly, define the set-valued $K : \ell^\infty \times \mathcal{Z}^+ \rightsquigarrow \mathcal{R}^n$ given by,

$$\begin{aligned} K(y,N) = \{ & \xi \in \mathcal{R}^n : \xi = (T_{xw}w + T_{xx_o}x_o)(N) \\ & \text{for some } (w, x_o) \in W(y,N) \}, \end{aligned}$$

i.e., $K(y,N)$ denotes the set of possible state-vectors at time N consistent with the measured data up to time N. Finally, define the set-valued $\tilde{K} : \mathcal{R} \rightsquigarrow \mathcal{R}^n$ by

$$\tilde{K}(v) = \{ \xi \in \mathcal{R}^n : v = C\xi + \nu \text{ for some } |\nu| \le 1 \}.$$

The set $\tilde{K}(v)$ represents the set of possible states based on a single measurement.

The following algorithm (see also [1, Section 20]) propagates the set of possible states in a recursive manner.

Algorithm 3.1 *Let* $y \in \ell^\infty$ *be a prescribed measurement trajectory.*

Initialization

$$K(y,0) = \tilde{K}(y(0)) \bigcap B_{\mathcal{R}^n}.$$

Propagation

$$K(y,N) = \tilde{K}(y(N)) \bigcap \left\{ \xi : \xi = A\tilde{\xi} + B\delta \text{ for some } \tilde{\xi} \in K(y, N-1), |\delta| \le 1 \right\}.$$

Note that all sets are constructed with a *causal* dependence on the measurement trajectory, y.

The following theorem describes a computational implementation of Algorithm 3.1.

Theorem 3.1 *In the framework of Algorithm 3.1,*

$$\tilde{K}(y(N)) = Set(\tilde{M}(N), \tilde{m}(N)),$$

where

$$\tilde{M}(N) = \begin{pmatrix} C \\ -C \end{pmatrix}, \quad \tilde{m}(N) = \begin{pmatrix} 1 + y(N) \\ 1 - y(N) \end{pmatrix},$$

and

$$K(y,N) = Set(M(N), m(N)),$$

where $(M(N), m(N))$ *belong to the Rack* $[\cdot]$ *of the matrices*

$$\begin{pmatrix} M(N-1)A^{-1} & -M(N-1)A^{-1}B \\ 0 & 1 \\ 0 & -1 \\ \tilde{M}(N) & 0 \end{pmatrix}, \begin{pmatrix} m(N-1) \\ 1 \\ 1 \\ \tilde{m}(N) \end{pmatrix}.$$

Proof The condition $x(N) \in \tilde{K}(y(N))$ is equivalent to

$$|y(N) - Cx(N)| \le 1,$$

which is equivalent to

$$\begin{pmatrix} C \\ -C \end{pmatrix} x(N) \le \begin{pmatrix} 1 + y(N) \\ 1 - y(N) \end{pmatrix},$$

which is the matrix description of $\tilde{K}(N)$.

Now according to Algorithm 3.1, the condition $x(N) \in K(y,N)$ is equivalently described by the two conditions $x(N) \in \tilde{K}(y(N))$ and

$$A^{-1}x(N) - A^{-1}B\delta \in K(y, N-1)$$

for some $|\delta| \leq 1$. Using that $K(y, N-1) = Set(M(N-1), m(N-1))$ leads to the equivalent statement,

$$\begin{pmatrix} M(N-1)A^{-1} & -M(N-1)A^{-1}B \\ 0 & 1 \\ 0 & -1 \\ \tilde{M}(N) & 0 \end{pmatrix} \begin{pmatrix} x(N) \\ \delta \end{pmatrix} \leq \begin{pmatrix} m(N-1) \\ 1 \\ 1 \\ \tilde{m}(N) \end{pmatrix},$$

for some δ. An application of the *Rack* $[\cdot]$ operator leads to the desired result. ∎

We see that the set of possible states forms a polytope described by a collection of inequalities. The computational burden of a real-time implementation amounts to the computation of the *Rack* $[\cdot]$ operator, which essentially requires the solution of several small linear programs to remove redundant constraints. The *Rack* $[\cdot]$ operator is a notational convenience for the Fourier-Motzkin algorithm described in [3]. For the sake of completeness, this is reviewed in Appendix. Since these sets may be described by several inequalities, the real-time applicability of these methods is questionable. This consideration has led to the construction of approximate simplified descriptions of $K(y, N)$, in particular through bounding ellipsoids. See [1, 5] and references contained therein for further discussion on these topics.

4 ℓ^∞ Induced-Norm Optimal Estimation

4.1 Problem Formulation and Main Result

In this section, we show that the set-valued observer in Section 3 can be used to provide optimal estimates in an induced-norm sense.

Define the scalar variable

$$z(k) = H(k)x(k).$$

As in Section 3, define T_{zw} and T_{zx_o} as the mappings from exogenous signals and initial conditions, respectively, to z.

We now define the our optimal estimation problem.

Definition 4.1 *An* **estimator** *is any causal (possibly nonlinear) mapping* $\Phi : \ell^\infty \to \ell^\infty$.

Definition 4.2 *The estimator* Φ^* *is* **pointwise optimal** *if for any other estimator,* Φ,

$$\sup_{(w,x_o)\in W(y,N)} \frac{|(T_{zw}w + T_{zx_o})(N) - (\Phi^* y)(N)|}{\|(w, x_o)\|}$$

$$\leq \sup_{(w,x_o)\in W(y,N)} \frac{|(T_{zw}w + T_{zx_o})(N) - (\Phi y)(N)|}{\|(w, x_o)\|},$$

for all $N \in \mathcal{Z}^+$ *and all possible measurement trajectories.*

The estimator Φ^* *is* **uniformly optimal** *if for any other estimator,* Φ,

$$\sup_{\substack{\|(w,x_o)\|\leq 1 \\ y = T_{yw}w + T_{yx_o}x_o}} \frac{|(T_{zw}w + T_{zx_o}x_o)(N) - (\Phi^* y)(N)|}{\|(w, x_o)\|} \leq \sup_{\substack{\|(w,x_o)\|\leq 1 \\ y = T_{yw}w + T_{yx_o}x_o}} \frac{|(T_{zw}w + T_{zx_o}x_o)(N) - (\Phi y)(N)|}{\|(w, x_o)\|},$$

for all $N \in \mathcal{Z}^+$.

Reference [7] considers the uniformly optimal estimation problem. In the case of zero-initial conditions and time-invariant dynamics, the uniformly optimal estimation problem can be solved as a standard ℓ_1 model-matching problem (cf., [2]). For non-zero initial conditions, the resulting model matching problem is time-varying. The resulting solution leads to a uniformly optimal estimator which is not recursive, i.e., the optimal estimate at time N requires storage of all measurements $\{y(0), \ldots, y(N)\}$. Reference [7] goes on to provide an approximately optimal estimator which is recursive after a fixed number of time-steps.

The following proposition summarizes the results of [7] needed here.

Proposition 4.1 ([7]) *There exists a uniformly optimal linear (time-varying) estimator,* Q. *Furthermore, the associated worst-case estimation error,* $\gamma(N)$, *defined by*

$$\gamma(N) = \sup_{\substack{\|(w,x_o)\|\leq 1 \\ y = T_{yw}w + T_{yx_o}x_o}} \frac{\|(T_{zw}w + T_{zx_o}x_o)(N) - (Qy)(N)\|}{\|(w, x_o)\|},$$

satisfies

$$\gamma(N) = \sup_{\substack{\|(w,x_o)\|\leq 1 \\ 0 = T_{yw}w + T_{yx_o}x_o}} \frac{\|(T_{zw}w + T_{zx_o}x_o)(N)\|}{\|(w, x_o)\|}.$$

Proposition 4.1 states that the cost of the uniformly optimal estimator (at any fixed N) is given by the worst-case estimation error incurred for the measurement trajectory $y = 0$.

The present estimation problem considers non-zero initial conditions *and* time-varying dynamics. We will show that the set-valued observer in Section 3 defines an pointwise optimal estimator. In addition to pointwise optimality being stronger than uniform optimality, an advantage is that the construction in Section 3 is *recursive*.

Definition 4.3 *Consider the set-valued observer of Algorithm 3.1. Define*

$$\underline{z}(y, N) = \min \{\zeta : \zeta \in K(y, N)\},$$

$$\overline{z}(y, N) = \max \{\zeta : \zeta \in K(y, N)\},$$

$$z_c(y, N) = \frac{\overline{z}(y, N) + \underline{z}(y, N)}{2}.$$

The **central estimator**, $\Phi_c : \ell^\infty \to \ell^\infty$, *is defined as*

$$(\Phi_c y)(N) = z_c(y, N).$$

Our main result is the following.

Theorem 4.1 *The central estimator, Φ_c, is pointwise optimal.*

4.2 Proof of Main Result

This section is devoted to the proof of Theorem 4.1.

Since we are interested in pointwise optimality, we will consider a single "experiment", i.e., a fixed measurement trajectory, y, and estimation time, N. This will simplify the presentation a great deal by dropping notational dependence on y and N throughout. Thus, for this *fixed* measurement trajectory, y, and estimation time, N, we will use the following shorthand notation:

- $\underline{z}, \overline{z}$, and z_c—rather than $\underline{z}(y, N)$, $\overline{z}(y, N)$, and $z_c(y, N)$,
- z_Q—rather than $(Qy)(N)$, and
- γ—rather than $\gamma(N)$,

where Q is the uniformly optimal estimator as in Proposition 4.1 and $\gamma(N)$ is the associated cost at time N.

Define $r : [\underline{z}, \overline{z}] \to \mathcal{R}^+$ by

$$r(\hat{z}) = \min \{\|(w, x_o)\| : \\ (w, x_o) \in W(y, N) \text{ and } \hat{z} = (T_{zw}w + T_{zx_o}x_o)(N).\}.$$

In other words, $r(\hat{z})$ is the size of the smallest exogenous signal/initial condition pair which can produce the measured output as well as the value $\hat{z}$.

Similarly, define $\phi : [\underline{z}, \overline{z}] \to \mathcal{R}$ by

$$\phi(\hat{z}) = \frac{\hat{z} - z_c}{r(\hat{z})}.$$

Then the estimation error associated with z_c can be expressed alternatively as

$$\sup_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi(\hat{z})|,$$

(compare to Definition 4.2).

Note that

$$r(\underline{z}) = r(\overline{z}) = 1.$$

Similarly,

$$\phi(\underline{z}) = -\phi(\overline{z}).$$

However, $\phi(\cdot)$ need not be a symmetric (odd) function. Furthermore we see r and ϕ can be derived from appropriate minimum distance problems, and are both *continuous* functions.

Claim 4.1 *The following inequality holds,*

$$\frac{\overline{z} - \underline{z}}{2} \leq \gamma.$$

In case of equality, $z_c = z_Q$.

Proof The uniformly optimal estimator satisfies

$$|\overline{z} - z_Q| \leq \gamma r(\overline{z}),$$

and

$$|\underline{z} - z_Q| \leq \gamma r(\underline{z}).$$

Since $r(\overline{z}) = r(\underline{z}) = 1$, this leads to

$$\overline{z} - \underline{z} \leq 2\gamma.$$

In case of equality, $z_Q = z_c$ is necessary. For example if $z_Q < z_c$,

$$\overline{z} - z_Q > \gamma r(\overline{z}),$$

which is a contradiction. ■

Claim 4.2 *Suppose $z_c \leq z_1 \leq \overline{z}$. Suppose $\phi(z_1) \leq \gamma$. Then $\phi(z_2) \geq \phi(z_1)$ for all $z_2 \geq z_1$.*

Proof Let (w_1, x_{o1}) produce z_1 with minimum norm, i.e., $(w_1, x_{o1}) \in W(y, N)$,

$$\|(w_1, x_{o1})\| = r(z_1),$$

and

$$z_1 = (T_{zw}w_1 + T_{zx_o}x_{o1})(N).$$

Let $(w_*, x_{o*}) \subset B_{\ell^\infty} \times B_{\mathcal{R}^n}$ correspond to the worst case exogenous signal/initial condition pair for the uniformly optimal observer as in Proposition 4.1. That is,

$$\gamma = \frac{|(T_{zw}w_* + T_{zx_o}x_{o*})(N)|}{\|(w_*, x_{o*})\|},$$

and

$$0 = (T_{yw}w_* + T_{yx_o}x_{o*})(N).$$

Without loss of generality, assume that

$$z_* = (T_{zw}w_* + T_{zx_o}x_{o*})(N) > 0.$$

One way to produce z_2 is through

$$(w_2, x_{o2}) = (w_1, x_{o1}) + h(w_*, x_{o*}),$$

where h is appropriately scaled so that

$$hz_* = z_2 - z_1.$$

By construction, (w_2, x_{o2}) consistent with the measured data. However, it may be that $\|(w_2, x_{o2})\| > 1$.

We now compare $\phi(z_2)$ and $\phi(z_1)$. First,

$$\phi(z_2) = \frac{z_2 - z_c}{r(z_2)} \geq \frac{z_2 - z_c}{\|(w_2, x_{o2})\|}$$
$$\geq \frac{z_2 - z_c}{\|(w_1, x_{o1})\| + h\,\|(w_*, x_{o*})\|}.$$

Thus proving the claim can be achieved by testing whether

$$\frac{z_2 - z_c}{\|(w_1, x_{o1})\| + h\,\|(w_*, x_{o*})\|} \geq \frac{z_1 - z_c}{\|(w_1, x_{o1})\|} = \phi(z_1).$$

Towards this end, we see that

$$\frac{z_2 - z_c}{\|(w_1, x_{o1})\| + h\,\|(w_*, x_{o*})\|} \geq \frac{z_1 - z_c}{\|(w_1, x_{o1})\|}$$

$$\Leftrightarrow$$

$$\big((z_2 - z_1) + (z_1 - z_c)\big)\,\|(w_1, x_{o1})\|$$
$$\geq (z_1 - z_c)\big(\|(w_1, x_{o1})\| + h\,\|(w_*, x_{o*})\|\big)$$

$$\Leftrightarrow$$

$$(z_2 - z_1)\,\|(w_1, x_{o1})\| \geq h(z_1 - z_c)\,\|(w_*, x_{o*})\|$$

$$\Leftrightarrow$$

$$(z_2 - z_1)\,\|(w_1, x_{o1})\| \geq \frac{z_2 - z_1}{z_*}(z_1 - z_c)\,\|(w_*, x_{o*})\|$$

$$\Leftrightarrow$$

$$\gamma = \frac{z_*}{\|(w_*, x_{o*})\|} \geq \frac{z_1 - z_c}{\|(w_1, x_{o1})\|} = \phi(z_1).$$

Using the hypothesis,

$$\phi(z_1) \leq \gamma$$

completes the proof. ∎

Claim 4.3 *Suppose $z_c \geq z_1 \geq \underline{z}$. Suppose $-\phi(z_1) \leq \gamma$. Then $\phi(z_2) \leq \phi(z_1)$ for all $z_2 \leq z_1$.*

Proof The proof is similar to the proof of Claim 4.2. ∎

Claim 4.4 *The function ϕ is monotonically non-decreasing over the interval $[\underline{z}, \overline{z}]$.*

Proof Claims 4.2–4.3 imply that ϕ is monotonic as long as $|\phi(\hat{z})| \leq \gamma$.

Note that $\phi(z_c) = 0$. Thus by continuity, ϕ is monotonic until $\phi(\hat{z}') = \gamma$ for some $\hat{z}' \in (\underline{z}, \overline{z})$. Assume that such a $\hat{z}'$ satisfies $\hat{z}' > z_c$. Similar arguments hold in case $\hat{z}' < z_c$. Since $\overline{z} \geq \hat{z}'$, Claim 4.2 implies $\phi(\overline{z}) \geq \gamma$, and hence $\phi(\underline{z}) \leq -\gamma$. Claim 4.1 then implies that actually

$$\phi(\overline{z}) = \gamma = -\phi(\underline{z}),$$

and $z_c = z_Q$. Since $z_c = z_Q$,

$$\phi(\hat{z}) = \frac{\hat{z} - z_Q}{r(\hat{z})} \leq \gamma.$$

Thus, if ever $\phi(\hat{z}') = \gamma$, then $\phi(\hat{z}) \leq \gamma$ for all $\hat{z} \in [\underline{z}, \overline{z}]$, which completes the proof. ∎

The proof of Claim 4.4 states that the function ϕ saturates at $\pm\gamma$ if it ever achieves these values. In this case, $z_c = z_Q$.

We can now show that z_c is the pointwise optimal estimate. The cost of an alternative estimate, z', may be expressed

$$\max_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi'(\hat{z})|,$$

where

$$\phi'(\hat{z}) = \frac{\hat{z} - z'}{r(\hat{z})}.$$

In case $z' < z_c$, then

$$\phi'(\overline{z}) > \phi(\overline{z}).$$

In case $z' > z_c$, then

$$\phi'(\underline{z}) < \phi(\underline{z}).$$

In either case,

$$\max_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi'(\hat{z})| > \max_{\hat{z} \in [\underline{z}, \overline{z}]} |\phi(\hat{z})|,$$

which completes the proof of Theorem 4.1.

5 Concluding Remarks

This paper has considered the guaranteed state-estimation problem for discrete-time linear time-varying systems. Based on an *a priori* model of initial conditions and exogenous signals, a set-valued observer was constructed which recursively computes the set of possible state vectors consistent with measured output data. It was shown that the centers of these sets correspond to the optimal state-estimate which minimizes the induced norm from exogenous signals/initial conditions to estimation error. The algorithms easily can be modified in the case of known initial conditions simply by changing the *a priori* assumptions.

The estimation problems considered here were for scalar-valued disturbances, noises, and estimates. However, the multivariable case requires only notational changes.

A Fourier-Motzkin Elimination Algorithm

For the sake of completeness, we review the Fourier-Motzkin elimination algorithm described in [3].

For $M_1 \in \mathcal{R}^{z\times n}$, $M_2 \in \mathcal{R}^z$, and $m \in \mathcal{R}^z$ consider the subset, S, of $\mathcal{R}^n$ defined by

$$S = \{x : M_1 x + M_2 w \leq m \text{ for some } w \in \mathcal{R}\}.$$

Define

$$\begin{aligned} &Rack\left[\begin{pmatrix} M_1 & M_2 \end{pmatrix}, m\right] \\ &\quad = \left\{(\tilde{M}, \tilde{m}) \in \mathcal{R}^{z\times n} \times \mathcal{R}^z : S = Set(\tilde{M}, \tilde{m})\right\}. \end{aligned}$$

Thus, $Rack\left[\begin{pmatrix} M_1 & M_2 \end{pmatrix}, m\right]$ is the set of matrices which give a direct characterization of S.

We now construct an element of $Rack\left[\begin{pmatrix} M_1 & M_2 \end{pmatrix}, m\right]$. In the following, $A(i,:)$ denotes the i^{th} row of the matrix A, and $a(i)$ denotes the i^{th} element of the column vector a. We begin by defining the following indices:

$$\begin{aligned} I_+ &= \{i : M_2(i) > 0\}, \\ I_- &= \{i : M_2(i) < 0\}, \\ I_0 &= \{i : M_2(i) = 0\}. \end{aligned}$$

Now suppose

$$M_1 x + M_2 m \leq m.$$

Then

$$w \leq \frac{m(i) - M_1(i,:)x}{M_2(i)}, \quad \forall i \in I_+,$$

and

$$w \geq \frac{m(i) - M_1(i,:)x}{M_2(i)}, \quad \forall i \in I_-.$$

Thus for $x \in S$, every upper-bound on w must be larger than every lower bound on w. A pairwise comparison shows that $x \in S$ if and only if,

$$\frac{m(i_-) - M_1(i_-,:)x}{M_2(i_-)} \leq \frac{m(i_+) - M_1(i_+,:)x}{M_2(i_+)}$$

for all $i_- \in I_-$ and $i_+ \in I_+$, as well as

$$M_1(i_0,:)x \leq m(i_0), \quad \forall i_0 \in I_0.$$

Thus, let $\tilde{M}$ be the matrix formed by the rows:

$$\tilde{M}(j,:) = \rho_{i_+ i_-} = M_2(i_-)M_1(i_+,:) - M_2(i_+)M_1(i_-,:),$$

where a new row (indexed by j) is formed for each $i_+ \in I_+$ and $i_- \in I_-$, and

$$\tilde{M}(j,:) = \rho_{i_0} = M_1(i_0,:),$$

where a new row is formed for each $i_0 \in I_0$. Similarly, let $\tilde{m}$ be the column vector formed by the associated elements

$$\tilde{m}(j) = m_{i_+ i_-} = m(i_+)M_2(i_-) - m(i_-)M_2(i_+),$$

$$\tilde{m}(j) = m_{i_0} = m(i_0),$$

Then $(\tilde{M}, \tilde{m}) \in Rack\left[\begin{pmatrix} M_1 & M_2 \end{pmatrix}, m\right]$.

In general, this procedure creates redundant constraints which can be removed by solving appropriate linear programs.

References

[1] F.L. Chernousko. *State Estimation for Dynamic Systems.* CRC Press, Inc., Boca Raton, FL, 1994.

[2] M.A. Dahleh and I.J Diaz-Bobillo. *Control of Uncertain Systems: A Linear Programming Approach.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1995.

[3] S.S. Keerthi and E.G. Gilbert. Computation of minimum-time feedback control laws for discrete-time systems with state-control constraints. *IEEE Transactions on Automatic Control,* **AC–32**:432–435, 1987.

[4] A.B. Kurzhanski and V.M Veliov, editors. *Modeling Techniques for Uncertain Systems.* Birkhäuser, Boston, 1993.

[5] M. Milanese and V. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica,* **27**:997–1009, 1991.

[6] K.M. Nagpal and P.P. Khargonekar. Filtering and smoothing in an $\mathcal{H}^\infty$ setting. *IEEE Transactions on Automatic Control,* **AC–36**:152–166, 1991.

[7] P. Voulgaris. On optimal ℓ_∞ to ℓ_∞ filtering. In *Proceedings of the American Control Conference*, 1993. to appear in *Automatica* 1995.