Minimax state estimation for linear discrete-time differential-algebraic equations

Sergiy M.Zhuk

Department of System Analysis and Decision Making Theory, Taras Shevchenko National University of Kyiv, Ukraine

Abstract

This paper presents a state estimation approach for an uncertain linear equation with a non-invertible operator in Hilbert space. The approach addresses linear equations with uncertain deterministic input and noise in the measurements, which belong to a given convex closed bounded set. A new notion of a *minimax observable subspace* is introduced. By means of the presented approach, new equations describing the dynamics of a minimax recursive estimator for discrete-time non-causal differential-algebraic equations (DAEs) are presented. For the case of regular DAEs it is proved that the estimator's equation coincides with the equation describing the seminal Kalman filter. The properties of the estimator are illustrated by a numerical example.

Key words: Robust estimation; Descriptor systems; Optimization under uncertainties; Set-membership estimation; Minimax
1991 MSC: 93E11 93E10 60G35

1 Introduction

The importance of models described by DAEs (or descriptor systems) in economics, demography, mechanics and engineering is well known [13]. Here, motivated by further applications to linear DAEs, we present a state estimation approach for linear deterministic models described by an abstract linear equation in a Hilbert space. Our approach is based on ideas underlying H_2/H_∞ filtering [4,3] and set-membership state estimation [7,14,16,12,15].

H_2-estimators like Kalman or Wiener filters [4,1] give estimations of the system state with minimum error variance. The H_2-estimation problem for linear time-variant DAEs was studied in [17] without restricting the DAE's matrices. The resulting algorithm requires the calculation of the so-called "3-block matrix pseudoinverse". In [10] the authors introduced explicit formulas for the 3-block matrix pseudoinverse and derived a recurrence filter, assuming a special structure for the DAE's matrices. A brief overview of steady-state H_2-estimators is presented in [8].

H_∞ estimators minimize a norm of the operator mapping unknown disturbances with finite energy to filtered errors [3]. We stress that the H_∞ estimator coincides with a certain Krein space H_2 filter [19]. The H_∞ filtering technique was applied to linear time-invariant DAEs with regular matrix pencils in [22].

A basic notion in the theory of set-membership state estimation is that of an a posteriori set or informational set. This notion has roots in control theory [4]. By definition, it is the set of all possible state vectors φ, that are consistent with a measured output y, provided that an uncertain input f and measurement error η belong to some bounded set $\mathscr{G}$. We will be interested in the case when the state $\varphi \in \mathcal{H}$ obeys an abstract linear equation $L\varphi = f$, provided $y = H\varphi + \eta$, $(f, \eta) \in \mathscr{G}$, where $\mathscr{G}$ is a bounded closed convex subset of an abstract Hilbert space. The problem is to find an estimation $\hat{\varphi}$ of φ with minimal worst-case error. This problem was previously considered in [16,14]. Due to [16] a vector $\hat{\varphi}$ is called a *linear minimax a-posteriori estimation* (or a *central algorithm* due to [14]) iff $\forall \ell \in \mathcal{H}$

$$\langle \ell, \hat{\varphi} \rangle = (\sup_{\mathscr{G}(y)} \langle \ell, \varphi \rangle + \inf_{\mathscr{G}(y)} \langle \ell, \varphi \rangle)/2, \tag{1}$$

provided that an a posteriori set $\mathscr{G}(y) := \{\varphi : (L\varphi, y - H\varphi) \in \mathscr{G}\}$ is a bounded convex subset of the Hilbert space $\mathcal{H}$. Note that if there exists φ_0 so that $L\varphi_0 = 0$, $H\varphi_0 = 0$, then $\sup_{\mathscr{G}(y)} \langle \ell, \varphi \rangle = +\infty$ for some ℓ. Thus, the above approach does not work if L is non-injective. In this paper we generalize the approach of [14,16] to linear equations with non-injective L. Futher generalization is presented in [25].

⋆ This paper was presented at IFAC Workshop CAO09, Jyvaskyla, Finland, May 6-9, 2009. Corresponding author S. Zhuk. Tel. +38050-52-59138. Fax +38044-57-56684.

Email address: `beetle@unicyb.kiev.ua` (Sergiy M.Zhuk).

The main contribution of this paper is a new notion of a *minimax observable subspace* $\mathcal{L}$ for the pair (L,H) (Definition 1). It is useful when one needs to evaluate *a priori* how far the estimation $\hat{\varphi}$ is from a "real" state φ in the worst case, provided $\hat{\varphi}$ is constructed from the measurements y. Due to Proposition 1, the *worst-case estimation error* is finite iff $\mathcal{L}=\mathcal{H}$; otherwise $\hat{\varphi}$ may be too far from a "real" state φ for some directions $\ell\in\mathcal{H}$, even for bounded f and η. In fact, given y, we can provide an estimation with finite worst-case error for the projection of φ onto $\mathcal{L}$ only. Thus $\mathcal{L}$ is an analog of the observable subspace [4, p.240] for the pair (L,H) in the context of set-membership state estimation.

The introduced notion allows the generalization of ideas from [14,16,21] to non-injective linear mappings, in particular for the case $L\varphi(t)=((F\varphi)_t-C(t)\varphi(t),F\varphi(t_0))$ with $F\in\mathbb{R}^{m\times n}$ which arise in the state estimation for linear continuous non-causal [1] DAEs [24]. As a consequence, one can apply the minimax framework, originally developed [7,5] for DAEs ($F=E$ in the linear case) with bounded uncertainties, to DAEs [25] with unbounded inputs (see example in Section 3).

In order to stress connections with H_∞ approach, we note that the minimax framework [7,5] incorporates the set-membership state estimation and H_∞ filtering for ODEs by application of dynamic programming [2] to the informational state $X(\tau)$: for linear ODE the worst-case estimation is set to be the Tchebysheff center of $X(\tau)$. Although we derive the estimation from the minimization of the worst-case error, as it is stated in Definition 1, our approach (for the ellipsoidal bounding set $\mathscr{G}$ and causal [2] DAEs) results in the same estimation and error as in [7]. Thus, the ℓ-minimax estimation gives a proper generalization of the recurrence algorithm from [7] to the case when $X(\tau)$ may be unbounded.

We illustrate the benefits of the new notion by introducing a minimax recursive estimator for discrete-time non-causal DAEs: it works for non-causal DAEs unlike [10,8,22,23] and for the regular case it coincides with one proposed in [10] (Corollary 1). In addition, the minimax observability subspace allows one to identify the observable (in the minimax sense) part of the state with respect to given measurements. Computing the index of non-causality, one can a priori check how good connections between observations and state are: models with zero index are fully observable while models with non-zero index have an unobservable part in the state.

This paper is organized as follows. In Section 2 we give definitions (Definition 1) of the minimax estimation, error and observable subspace for abstract linear equations and we construct the estimation for a convex bounded $\mathscr{G}$, in particular for an ellipsoidal $\mathscr{G}$ (Proposition 1). In Section 3 we introduce the minimax observable subspace and index of non-causality for DAEs in discrete time (Definition 2) and we derive the minimax estimator (Theorem 1). Also we discuss connections to H_2/H_∞ framework (Corollary 1) and present an example.

Notation. *Linear mappings:* $\langle\cdot,\cdot\rangle$ denotes the inner product; $\mathscr{L}(\mathcal{H}_1,\mathcal{H}_2)$ denotes the space of all bounded linear mappings from $\mathcal{H}_1$ to $\mathcal{H}_2$, $\mathscr{L}(\mathcal{H}):=\mathscr{L}(\mathcal{H},\mathcal{H})$; $\mathbb{1}_{\mathcal{H}}$ is the identity mapping in $\mathcal{H}$; $\mathscr{D}(L)$, $R(L)$, $N(L)$ denote, respectively, the domain, range, and null-space of a linear mapping $L:\mathscr{D}(L)\mapsto R(L)$; $L^*:\mathcal{H}\to\mathcal{H}$ is the adjoint of L; F' denotes the transpose of F; F^+ is the pseudoinverse of F; E is the identity matrix; $\mathrm{diag}(A_1\dots A_n)$ denotes a diagonal matrix with A_i, $i=\overline{1,n}$ on its diagonal; $\{x_s\}_1^n:=(x_1,\dots,x_n)$ is an element of $\mathcal{H}_1\times\dots\times\mathcal{H}_n$, $0_{mn}\in\mathbb{R}^{m\times n}$ denotes the $m\times n$-zero matrix.
Functionals: $I_1(x):=\langle Q_1Lx,Lx\rangle+\langle Q_2Hx,Hx\rangle$, $I(x):=\langle Q_1Lx,Lx\rangle+\langle Q_2(y-Hx),y-Hx\rangle$ with positive definite self-adjoint $Q_1\in\mathscr{L}(\mathcal{F})$ and $Q_2\in\mathscr{L}(\mathcal{Y})$; $c(G,x):=\sup\{\langle x,y\rangle,y\in G\}$ is a support function of G; $\gamma_\pm:=\frac{1}{2}(c(\mathscr{G}(y),\ell)\pm c(\mathscr{G}(y),-\ell))$; $\|x\|_S^2=\langle Sx,x\rangle$.
Sets: $\mathscr{G}^\beta(0):=\{x:I_1(x)\le\beta\}$, $\mathscr{G}(0):=\{x:I_1(x)\le1\}$; $\mathrm{dom}f:=\{x:f(x)<\infty\}$ is an effective domain of f; $\mathrm{Argmin}_xf:=\{x:f(x)=\min_xf\}$ is the set of global minima of f; $\overline{G}$ is the closure of a set G.

2 Linear minimax estimation problem in a Hilbert space

Let vector y be observed in the form of

$$y=H\varphi+\eta \tag{2}$$

where φ obeys the equation

$$L\varphi=f \tag{3}$$

We assume that $L:\mathscr{D}(L)\mapsto\mathcal{F}$ is a closed linear operator [4, p.63], $\mathscr{D}(L)\subset\mathcal{H}$ is a linear set, $\overline{\mathscr{D}(L)}=\mathcal{H}$, $H\in\mathscr{L}(\mathcal{H},\mathcal{Y})$. Also we assume that (f,η) is an unknown element of a convex bounded closed set $\mathscr{G}\subset\mathcal{F}\times\mathcal{Y}$, $\mathcal{H},\mathcal{F},\mathcal{Y}$ are Hilbert spaces.

Definition 1 *Let* $\mathcal{L}:=\{\ell\in\mathcal{H}:\hat{\rho}(\ell)<\infty\}$ *with*

$$\hat{\rho}(\ell):=\inf_{\varphi\in\mathscr{G}(y)}\rho(\ell,\varphi),\quad \rho(\ell,\varphi):=\sup_{\psi\in\mathscr{G}(y)}|\langle\ell,\varphi-\psi\rangle|$$

The set $\mathcal{L}$ *is called a minimax observable subspace for the pair* (L,H). *A vector* $\hat{\varphi}\in\mathscr{G}(y)$ *is called a minimax a posteriori estimation in the direction* [3] ℓ *(ℓ-estimation) if* $\rho(\ell,\hat{\varphi})=\hat{\rho}(\ell)$. *The number* $\hat{\rho}(\ell)$ *is called a minimax a posteriori error in the direction* ℓ *(ℓ-error).*

[1] DAE is said to be non-causal if the corresponding initial-value problem has more than one solution.

[2] Note that the dynamic programming was previously applied to causal DAEs in [6] in order to construct a regulator in LQ-control problems with DAE constraints.

[3] The subspace $l:=\{\alpha\ell:\alpha\in\mathbb{R}\}$, assigned with $\ell\in\mathcal{H}$, defines some direction in $\mathcal{H}$. To estimate φ in the direction ℓ means to estimate the projection $\langle\ell,\varphi\rangle\ell$ of φ onto l.

Our aim here is, given y, to construct the ℓ-estimation $\hat{\varphi}$ of the state φ, ℓ-error $\hat{\rho}(\ell)$ and minimax observable subspace $\mathcal{L}$, provided $\ell \in \mathcal{L}$. Note that $\hat{\rho}(\ell) = +\infty$ if $\ell \notin \mathcal{L}$ so that any $\psi \in \mathcal{H}$ is a ℓ-estimation by Definition 1.

Proposition 1 *Let $\mathscr{G}$ be a convex closed bounded subset of $\mathcal{F} \times \mathcal{Y}$. Then $\ell \in \mathcal{L}$ iff $\ell, -\ell \in \operatorname{dom} c(\mathscr{G}(y), \cdot)$. If $\ell \in \mathcal{L}$ then the ℓ-estimation $\hat{\varphi}$ along with ℓ-error $\hat{\rho}(\ell)$ obey*

$$\langle \ell, \hat{\varphi} \rangle = \gamma_-, \hat{\rho}(\ell) = \gamma_+ \tag{4}$$

Define $T : \mathscr{D}(T) \to \mathcal{H}$ by the rule $T(z,u) := L^* z + H^* u$ with $\mathscr{D}(T) := \mathscr{D}(L^*) \times \mathcal{Y}$ and let

$$\mathscr{G} = \{(f, \eta) : \langle Q_1 f, f \rangle + \langle Q_2 \eta, \eta \rangle \leq 1\}$$

If $R(T) = \overline{R(T)}$ then $\hat{x} \in \operatorname{Argmin}_x I$ is the ℓ-estimation, $\mathcal{L} = \operatorname{dom} c(\mathscr{G}(0), \cdot) = R(T)$ and

$$\hat{\rho}(\ell) = (1 - I(\hat{x}))^{\frac{1}{2}} c(\mathscr{G}(0), \ell). \tag{5}$$

The worst-case estimation error for any direction is

$$\begin{aligned} &\sup_{x \in \mathscr{G}(y)} \|\hat{x} - x\| = \inf_{\varphi \in \mathscr{G}(y)} \sup_{x \in \mathscr{G}(y)} \|\varphi - x\| = \\ &(1 - I(\hat{x}))^{\frac{1}{2}} \sup_{\|\ell\|=1} c(\mathscr{G}(0), \ell) = \sup_{\|\ell\|=1} \hat{\rho}(\ell) < +\infty \end{aligned} \tag{6}$$

If $\mathcal{L} = \mathcal{H}$ then (6) is finite.

PROOF. Let $\ell(\mathscr{G}(y)) = \{\langle \ell, \psi \rangle, \psi \in \mathscr{G}(y)\}$. Since $\mathscr{G}(y)$ is convex (due to convexity of $\mathscr{D}(L)$ and $\mathscr{G}$) and $x \mapsto \langle \ell, x \rangle$ is continuous, it follows that $\overline{\ell(\mathscr{G}(y))}$ is connected. Noting that $\inf_{\mathscr{G}(y)} \langle \ell, \psi \rangle = -c(\mathscr{G}(y), -\ell)$, we see

$$\overline{\ell(\mathscr{G}(y))} = [-c(\mathscr{G}(y), -\ell), c(\mathscr{G}(y), \ell)] \subset \mathbb{R}^1$$

Thus $\rho(\ell, \varphi) = +\infty$ if $\ell, -\ell \notin \operatorname{dom} c(\mathscr{G}(y), \cdot)$. Otherwise $\overline{\ell(\mathscr{G}(y))}$ is bounded, implying $\ell \in \mathcal{L}$. Hence, $\ell \in \mathcal{L}$ iff $\ell, -\ell \in \operatorname{dom} c(\mathscr{G}(y), \cdot)$.
Let $\ell \in \mathcal{L}$, $\varphi \in \mathscr{G}(y)$. Since $\overline{\ell(\mathscr{G}(y))}$ is connected, there exists $\varphi_* \in \mathscr{G}(y)$ so that $\langle \ell, \varphi_* \rangle = \gamma_-$ is the central point of $\overline{\ell(\mathscr{G}(y))}$. The worst-case distance $\rho(\ell, \varphi)$ is equal to the sum of the distance $|\langle \ell, \varphi - \varphi^* \rangle|$ between $\langle \ell, \varphi \rangle$ and the central point γ_- and the distance γ_+ between one of the boundary points of $\overline{\ell(\mathscr{G}(y))}$ and γ_-. Therefore, γ_- has the minimal worst-case distance γ_+. Hence, $\varphi_* = \hat{\varphi}$ due to Definition 1, which implies (4).
We proceed with the ellipsoidal $\mathscr{G}$. Let $Q_1 = \mathbb{1}_{\mathcal{F}}, Q_2 = \mathbb{1}_{\mathcal{Y}}$ for a simplicity. Due to [11, Sec 5.§3] $R(T) = \overline{R(T)}$ implies $R(T^*) = \overline{R(T^*)}$. Thus [4, p.14,Cor.1.4.3], there exists $\hat{x} \in \mathscr{D}(L)$ so that $T^* \hat{x}$ is the projection of $(0, y)$ onto $R(T^*) = \{(Lx, Hx), x \in \mathscr{D}(L)\}$, implying $\hat{x} \in \operatorname{Argmin}_x I$, and

$$\langle y - H\hat{x}, Hx \rangle = \langle L\hat{x}, Lx \rangle, \forall x \in \mathscr{D}(L)$$

Noting this, one easily derives [4] $I(\hat{x} - x) = I_1(x) + I(\hat{x})$ for all $x \in \mathscr{D}(L)$. Having it in mind and noting that $\mathscr{G}(y) = \{\varphi : I(\varphi) \leqslant 1\}$ and $\hat{x} \in \mathscr{G}(y)$, one derives

$$\hat{x} + \mathscr{G}^\beta(0) = \mathscr{G}(y) \tag{7}$$

where $\beta := 1 - I(\hat{x})$. (7) implies [18, p.113]

$$c(\mathscr{G}(y), \ell) = \langle \ell, \hat{x} \rangle + c(\mathscr{G}^\beta(0), \ell), \forall \ell \in \mathcal{H} \tag{8}$$

The definition of γ_-, (8) and $c(\mathscr{G}^\beta(0), \ell) = c(\mathscr{G}^\beta(0), -\ell)$ imply $\gamma_- = \langle \ell, \hat{x} \rangle$. Due to (4), $\hat{x}$ is the ℓ-estimation.
Let us prove (5). If $x \in \mathscr{G}^\beta(0)$ then $I_1(\beta^{-\frac{1}{2}} x) \leq 1$. Thus $\beta^{-\frac{1}{2}} \mathscr{G}^\beta(0) \subset \mathscr{G}(0)$ implying $\mathscr{G}^\beta(0) \subset \beta^{\frac{1}{2}} \mathscr{G}(0)$. If $x \in \mathscr{G}(0)$ then $I_1(\beta^{\frac{1}{2}} x) = \beta I_1(x) \leq \beta \Rightarrow \beta^{\frac{1}{2}} \mathscr{G}(0) \subset \mathscr{G}^\beta(0)$. Therefore

$$\mathscr{G}^\beta(0) = \beta^{\frac{1}{2}} \mathscr{G}(0) \Rightarrow c(\mathscr{G}^\beta(0), \ell) = \beta^{\frac{1}{2}} c(\mathscr{G}(0), \ell) \tag{9}$$

Now (4) and (8) imply (5). Hence, $\mathcal{L} = \operatorname{dom} c(\mathscr{G}(0), \cdot)$ due to Definition 1.
Let us prove $R(T) = \mathcal{L}$. If $\ell \in R(T)$ then $\ell = L^* z + H^* u$ for some $z \in \mathscr{D}(L^*)$, $u \in \mathcal{Y}$ and we get $\forall \varphi \in \mathscr{G}(0)$ by Cauchy inequality [4, p.4]

$$\langle \ell, \varphi \rangle = \langle z, L\varphi \rangle + \langle u, H\varphi \rangle \leq \|z\|^2 + \|u\|^2 < +\infty$$

so that $R(T) \subset \operatorname{dom} c(\mathscr{G}(0), \cdot) = \mathcal{L}$. If $\ell \notin R(T)$ then $\langle \ell, x \rangle > 0$ for some $x \in N(T^*)$ as $\mathcal{H} = \overline{R(T)} \oplus N(T^*)$. Noting that $\mathscr{G}(0) = \{x : \|T^* x\|^2 \leq 1\}$ we derive $c(\mathscr{G}(0), \ell) \geq \sup\{\langle \ell, x \rangle : T^* x = 0\} = +\infty$. Hence, $\mathcal{L} \subset R(T)$.
Let us prove (6). Set $\hat{\alpha} := \inf_{\varphi \in \mathscr{G}(y)} \sup_{x \in \mathscr{G}(y)} \|\varphi - x\|$. Using Definition 1, one derives [5]

$$\begin{aligned} \hat{\alpha} = &\inf_{\varphi \in \mathscr{G}(y)} \sup_{x \in \mathscr{G}(y)} \sup_{\|\ell\|=1} |\langle \ell, \varphi - x \rangle| \geq \\ &\sup_{\|\ell\|=1} \inf_{\varphi \in \mathscr{G}(y)} \sup_{x \in \mathscr{G}(y)} |\langle \ell, \varphi - x \rangle| = \sup_{\|\ell\|=1} \hat{\rho}(\ell) = \\ &\sup_{\|\ell\|=1} \sup_{x \in \mathscr{G}(y)} |\langle \ell, \hat{x} - x \rangle| = \sup_{x \in \mathscr{G}(y)} \|\hat{x} - x\| \geq \hat{\alpha} \end{aligned}$$

Now assume $\mathcal{L} = \mathcal{H}$. Since $\operatorname{dom} c(\mathscr{G}(0), \cdot) = \mathcal{L}$, it follows $c(\mathscr{G}(0), \cdot)$ is finite in $\mathcal{H}$ and therefore [9, §2.3] continuous. As a consequence, (6) is finite.

[4] $I(\hat{x} - x) = I_1(x) + I(\hat{x}) - 2\langle L\hat{x}, Lx \rangle + 2\langle y - H\hat{x}, Hx \rangle$
[5] Note [4, p.42] that $\|\varphi\| = \sup_{\|\ell\|=1} \langle \ell, \varphi \rangle$ and [4, p.55] $\inf_x \sup_y F(x,y) \geq \sup_y \inf_x F(x,y)$ for convex-concave F.

3 ℓ-estimation for non-causal DAEs

Consider the model

$$F_{k+1}x_{k+1} - C_k x_k = f_{k+1}, F_0 x_0 = f_0, \quad (10)$$
$$y_k = H_k x_k + g_k, k = 0, 1, \dots \quad (11)$$

where $F_k, C_k \in \mathbb{R}^{m\times n}$, $H_k \in \mathbb{R}^{p\times n}$, $x_k \in \mathbb{R}^n$ is a state, $f_k \in \mathbb{R}^m$ is an input and $y_k, g_k \in \mathbb{R}^p$ represent an output and the output's noise respectively. In what follows we assume that an initial state x_0 belongs to the affine set $\{x : F_0 x = f_0\}$. We define $\xi_\tau = (\{f_s\}_0^\tau, \{g_s\}_0^\tau)$ and assume

$$\xi_\tau \in \mathscr{G} = \{\xi_\tau : \sum_0^\tau \langle S_i f_i, f_i\rangle + \langle R_i g_i, g_i\rangle \leqslant 1\} \quad (12)$$

where $S_k \in \mathbb{R}^{m\times m}$ and $R_k \in \mathbb{R}^{p\times p}$ are positive definite self-adjoint matrices.
Suppose we observe $y_1^*, \dots, y_\tau^*$, provided that y_k^* is derived from (11) with $g_k = g_k^*$ and $x_k = x_k^*$, which obeys (10) with $f_k = f_k^*$, and $(\{f_s^*\}_0^\tau, \{g_s^*\}_0^\tau) \in \mathscr{G}$. Denote by $X(\tau)$ the set of all possible states x_τ of (10) consistent with measurements $y_1^*, \dots, y_\tau^*$ and uncertainty description (12).

Definition 2 *We say that $\overline{x}_\tau$ is a ℓ-estimation of the state x_τ^* in the direction $\ell \in \mathbb{R}^n$ iff*

$$\sup_{x\in X(\tau)} |\langle \ell, x - \overline{x}_\tau\rangle| = \hat\rho(\ell,\tau) := \inf_{z\in X(\tau)} \sup_{x\in X(\tau)} |\langle \ell, x - z\rangle|$$

$\hat\rho(\ell,\tau)$ is said to be an ℓ-error. A minimax observable subspace at the instant τ for the model (10)-(11) is defined by $\mathcal{L}(\tau) = \{\ell : \hat\rho(\ell,\tau) < \infty\}$. $I_\tau := n - \dim\mathcal{L}(\tau)$ is called an index of non-causality of the model (10)-(11).

Theorem 1 *Define $\hat\beta_\tau := 1 - \alpha_\tau + \langle P_\tau \hat x_\tau, \hat x_\tau\rangle$ and $\hat x_\tau = P_\tau^+ r_\tau$ with $r_0 = H_0' R_0 y_0$, $\alpha_0 = \langle R_0 y_0, y_0\rangle$,*

$$P_k = H_k' R_k H_k + F_k'[S_{k-1} - S_{k-1}C_{k-1}B_{k-1}^+ C_{k-1}' S_{k-1}]F_k,$$
$$P_0 = F_0' S_0 F_0 + H_0' R_0 H_0, B_k = P_k + C_k' S_k C_k,$$
$$\alpha_k = \alpha_{k-1} + \langle R_k y_k, y_k\rangle - \langle B_{k-1}^+ r_{k-1}, r_{k-1}\rangle,$$
$$r_k = F_k' S_{k-1} C_{k-1} B_{k-1}^+ r_{k-1} + H_k' R_k y_k,$$

Then $\hat x_\tau$ is the ℓ-estimation of x_τ^*, $\mathcal{L}(\tau) = \{\ell : P_\tau^+ P_\tau \ell = \ell\}$, $\hat\rho(\ell,\tau) = \hat\beta_\tau^{\frac12}\langle P_\tau^+\ell,\ell\rangle^{\frac12}$ and [6]

$$X(\tau) = \hat x_\tau + \hat\beta_\tau^{\frac12}\tilde X(\tau), \tilde X(\tau) := \{x : \langle P_\tau x, x\rangle \le 1\} \quad (13)$$

[6] Note that $\tau \mapsto X(\tau)$ represents the a posteriori set-valued observer [20]

PROOF. In order to apply Proposition 1, we rewrite (10)-(11) in the operator form: set $\mathcal{H} = (\mathbb{R}^n)^{\tau+1}$, $\mathcal{Y} = (\mathbb{R}^p)^{\tau+1}$, $\mathcal{F} = (\mathbb{R}^m)^{\tau+1}$ and $\varphi^* = \{x_s^*\}_0^\tau$, $y^* = \{y_s^*\}_0^\tau$, $f^* = \{f_s^*\}_0^\tau$, $\eta^* = \{g_s^*\}_0^\tau$,

$$L = \begin{pmatrix} F_0 & 0_{mn} & 0_{mn} & \dots & 0_{mn} & 0_{mn} \\ -C_0 & F_1 & 0_{mn} & \dots & 0_{mn} & 0_{mn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{mn} & 0_{mn} & 0_{mn} & \dots & -C_{\tau-1} & F_\tau \end{pmatrix}, H = \begin{pmatrix} H_0 & \dots & 0_{pn} \\ \dots & \dots & \dots \\ 0_{pn} & \dots & H_\tau \end{pmatrix}$$

Define $\mathcal{P}_\tau = \begin{pmatrix} 0_{nn}, & \dots, & 0_{nn}, E \end{pmatrix}$ and rewrite (12) with $Q_1 = \operatorname{diag}(S_0, \dots, S_\tau)$, $Q_2 = \operatorname{diag}(R_0, \dots, R_\tau)$ as

$$\xi_\tau = (f,\eta) \in \mathscr{G} = \{(f,\eta) : \langle Q_1 f, f\rangle + \langle Q_2\eta,\eta\rangle \le 1\}.$$

It is clear that y^*, H, η^*, L, f^*, φ^*, defined as above, satisfy (2)-(3) and $(f^*,\eta^*) \in \mathscr{G}$. Let $\mathscr{G}(y^*)$ denote the a posteriori set generated by y^*. Then $X(\tau) = \mathcal{P}_\tau(\mathscr{G}(y^*))$ by definition. Thus

$$\hat\rho(\ell,\tau) = \inf_{\varphi\in\mathscr{G}(y^*)} \sup_{\psi\in\mathscr{G}(y^*)} |\langle \ell, \mathcal{P}_\tau(\varphi - \psi)| = \hat\rho(l) \quad (14)$$

with $l = \mathcal{P}_\tau'\ell$. Hence, $\overline{x}_\tau = \mathcal{P}_\tau\hat\varphi$, where $\hat\varphi$ is the l-minimax estimation of the state φ^* of (3) in the sense of Definition 1. Proposition 1 implies $\hat\varphi = \hat x$ and $\hat\rho(l) = \beta^{\frac12} c(\mathscr{G}(0), l)$. Let us prove (13). (7) and (9) imply $\mathcal{P}_\tau(\mathscr{G}(y^*)) = \mathcal{P}_\tau\hat x + \beta^{\frac12}\mathcal{P}_\tau(\mathscr{G}(0))$. Therefore, $X(\tau) = \mathcal{P}_\tau\hat x + \beta^{\frac12}\mathcal{P}_\tau(\mathscr{G}(0))$. Now, let us prove $\hat x_\tau = \mathcal{P}_\tau\hat x$ by the direct calculation. Define

$$V_\tau(x_0, \dots, x_\tau) := \Phi(x_0) + \sum_{s=0}^{\tau-1}\Phi_s(x_s, x_{s+1})$$

with $\Phi_k(x,p) := \|F_{k+1}p - C_k x\|^2_{S_{k+1}} + \|y_{k+1} - H_{k+1}p\|^2_{R_{k+1}}$, $\Phi(x) := \|F_0 x\|^2_{S_0} + \|y_0 - H_0 x\|^2_{R_0}$, and

$$\mathcal{B}_\tau(p) := \min_{x_0\dots x_{\tau-1}} V_\tau(x_0 \dots x_{\tau-1}, p), \mathcal{B}_0(p) := \Phi(p)$$

Lemma 1. *Let $p \in \mathbb{R}^n$. There exists $(\tilde x_1 \dots \tilde x_{\tau-1}) \in (R^n)^{\tau-1}$ so that $V_\tau(\tilde x_1 \dots \tilde x_{\tau-1}, p) = \mathcal{B}_\tau(p)$ and*

$$\mathcal{B}_k(p) = \langle P_k p, p\rangle - 2\langle r_k, p\rangle + \alpha_k, P_k \ge 0 \quad (15)$$

Lemma 1 implies $\mathcal{B}_\tau$ is a quadratic and non-negative function. Therefore $\operatorname{Argmin}\mathcal{B}_\tau = \{x : P_\tau x = r_\tau\} \ne \varnothing$. This and $I(\varphi) = V_\tau(x_0, \dots, x_\tau)$ imply

$$\mathcal{B}_\tau(\hat x_\tau) = \min\mathcal{B}_\tau = \min V_\tau = \min I$$

Defining $\hat x := (\tilde x_1 \dots \tilde x_{\tau-1}, \hat x_\tau)$ with $\tilde x_k$ taken as in Lemma 1 for $p = \hat x_\tau$, we obtain $I(\hat x) = \min I$ and $\hat x_\tau = \mathcal{P}_\tau\hat x$. Therefore, $\hat x_\tau$ is a ℓ-estimation.
We note $\min I = \mathcal{B}_\tau(\hat x_\tau) = 1 - \hat\beta_\tau$. Thus $\beta = \hat\beta_\tau$ by definition.
Let us prove $\mathcal{P}_\tau(\mathscr{G}(0)) = \tilde X(\tau)$. Since $\mathscr{G}(0)$ does not depend on y^* and $\mathscr{G}(0) = \mathscr{G}(y^*)$ provided $y^* = 0$,

we can calculate $\mathcal{P}_\tau(\mathscr{G}(0))$ assuming $y^* = 0$. In this case $I = I_1 = V_\tau$, $\hat{x}_\tau = 0$ and $\mathcal{B}_\tau(x) = \langle P_\tau x, x\rangle$ so that $x \in \tilde{X}(\tau) \Leftrightarrow \mathcal{B}_\tau(x) \le 1$. If $x \in \mathcal{P}_\tau(\mathscr{G}(0))$ then, by definition, there exist $x_1 \dots x_{\tau-1}$ so that $V_\tau(x_1 \dots x_{\tau-1}, x) \le 1$, implying $\mathcal{B}_\tau(x) \le 1$. Now, let $\mathcal{B}_\tau(x) \le 1$. Then $V_\tau(\tilde{x}_1 \dots \tilde{x}_{\tau-1}, x) = \mathcal{B}_\tau(x)$ due to Lemma 1 and thus $x \in \mathcal{P}_\tau(\mathscr{G}(0))$ by definition.
Formulae (5), (14) and $\mathcal{P}_\tau(\mathscr{G}(0)) = \tilde{X}(\tau)$ imply $\hat{\rho}(\ell, \tau) = \hat{\beta}_\tau^{\frac{1}{2}} c(\tilde{X}(\tau), \ell) = \hat{\beta}_\tau^{\frac{1}{2}} \langle P_\tau^+ \ell, \ell\rangle^{\frac{1}{2}}$ for $\ell \in \mathcal{L}(\tau)$ and $\mathcal{L}(\tau) = \operatorname{dom} c(\tilde{X}(\tau), \cdot) = \{\ell : P_\tau^+ P_\tau \ell = \ell\}$. Details of calculation of $c(\tilde{X}(\tau), \cdot)$ are given in [18, p.108]. This completes the proof.

Proof of Lemma 1. We shall apply the dynamic programming [2]. Since V_τ is additive, it follows that

$$\mathcal{B}_\tau(p) = \min_{x_{\tau-1}} \{\Phi_{\tau-1}(x_{\tau-1}, p) + \mathcal{B}_{\tau-1}(x_{\tau-1})\} \tag{16}$$

V_τ is convex and non-negative by definition. Thus $\mathcal{B}_\tau$ is non-negative and convex for any $\tau \in \mathbb{N}$. Convexity is implied by the definition of $\mathcal{B}_\tau$ as for any convex function $(x, y) \mapsto f(x, y)$ the function $y \mapsto \min_x f(x, y)$ is convex [18, p.38].

Let us prove (15) by induction. (15) holds for $\mathcal{B}_0$ and P_0. We shall derive (15) for $\mathcal{B}_{k+1}$, P_{k+1}, assuming it holds for $\mathcal{B}_k$, P_k. Define

$$\Xi_k(x, p) := \Phi_k(x, p) + \langle P_k x, x\rangle - 2\langle r_k, x\rangle + \alpha_k \tag{17}$$

Then $\mathcal{B}_{k+1}(p) = \min_{x_k} \Xi_k(x_k, p)$ due to (16). Combining $P_k \ge 0$ with definition of Φ_k, we derive[7] $x \mapsto \Xi_k(x, p)$ is a convex quadratic function for any p. This and $\Xi_k(x_k, p) \ge B_{k+1}(p) \ge 0$ imply [18, p.268] $\operatorname{Argmin}_x \Xi_k(x, p) \ne \varnothing$. On the other hand [18, T.27.4] $x \in \operatorname{Argmin}_x \Xi_k(x, p)$ iff $\nabla_x \Xi_k(x, p) = 0$. Finally, we obtain $\operatorname{Argmin}_x \Xi_k(x, p) \ne \varnothing$ and

$$\operatorname{Argmin}_x \Xi_k(x, p) = \{x : B_k x = C_k' S_k F_{k+1} p + r_k\} \tag{18}$$

If we set $q_k = B_k^+(C_k' S_k F_{k+1} p + r_k)$ then $q_k \in \operatorname{Argmin}_x \Xi_k(x, p)$ due to [1]. Now, it is sufficient to calculate $\Xi_k(q_k, p)$ in order to see that (15) holds for $\mathcal{B}_{k+1}$ and P_{k+1}. Assertion $P_{k+1} \ge 0$ holds since $\mathcal{B}_{k+1}$ is convex. To conclude the proof, let us define $\tilde{x}_\tau = p$ and $\tilde{x}_k \in \operatorname{Argmin}_x \Xi_k(x, \tilde{x}_{k+1})$, $k = \overline{1, \tau-1}$. Then $V_\tau(\tilde{x}_1 \dots \tilde{x}_{\tau-1}, p) = \mathcal{B}_\tau(p)$ due to (16)-(18).

3.1. Example. Consider a system $p_{k+1} = A_k p_k + v_k, p_0 = v$ and assume $y_k = \tilde{H}_k p_k + g_k$ provided $p_k \in \mathbb{R}^n$ and $(v, g_0, \dots, g_\tau)$ belong to some ellipsoid. Now, given $y_1^*, \dots, y_\tau^*$, one needs to build the worst-case estimation of p_τ^*. We cannot apply directly standard minimax framework [16,7,5] in this case as we do not have any information about the bounding set for $(v_0, \dots, v_\tau)$. Instead, we apply the approach[8], proposed above. Define $F_k := (E, 0)$, $C_k := (A_k, E)$, $H_k := (\tilde{H}_k, 0)$ and $x_k^* := (p_k^*, v_k^*)$. Then x_k^* verifies (10)-(11) with $f_0 = v$, $f_k = 0$ and $g_k = g_k^*$, $k = \overline{1, \tau}$. Therefore, the original problem may be reformulated as: given $y_1^*, \dots, y_\tau^*$, one needs to build the $l = (\ell, 0)$-minimax estimation of x_τ^*. Of course, the estimation of x_τ in the direction $l = (0, \ell)$ has an infinite minimax error for any ℓ. But this is natural as $\langle l, x_\tau\rangle = \langle v_\tau, \ell\rangle$ and $(v_1, \dots, v_\tau)$ is unknown.
In what follows we present a numerical example. Let $p_k \in \mathbb{R}^2$, $\tilde{H}_k = \begin{pmatrix} 1 & 0 \end{pmatrix}$, $A_k = \begin{pmatrix} \frac{1}{10} & -\frac{1}{5} \\ \frac{7}{25} & -\frac{1}{10} \end{pmatrix}$ and $v_k = (-\frac{k\sin(k)}{10}, \frac{k\sin(k)}{10})$. We have generated p_k^* with $v = (\frac{1}{10}, \frac{1}{10})$ and y_k^* with $g_k = \frac{2\sin(k)}{k+1}$, $R_k = \frac{k}{k+1}$, $S_k = \operatorname{diag}(1, 1)$, $k = \overline{0, 50}$. The results are displayed on Fig. 1.

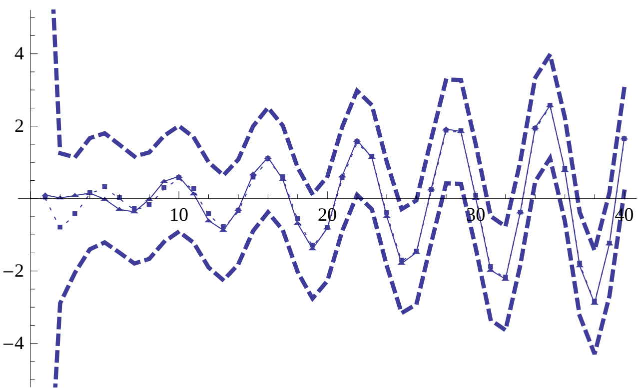

Fig. 1. The dashed line corresponds to the real values of $\langle \ell, p_k^*\rangle$ with $\ell = (0, 1)$, $k = \overline{1, 40}$; the solid line corresponds to the ℓ-estimation $\langle \ell, \hat{x}_k\rangle$; the bold dashed lines represent dynamics of the boundary points of the segment $\ell(\tilde{X}(\tau))$. Note that the trajectory of the estimation is centered with respect to "the bounds" – bold dashed lines.

3.2. Minimax estimator and H_2/H_∞ filters. In [5] a connection between set-membership state estimation and H_∞ approach is described for linear causal DAEs. The authors note that the notion of informational state ($X(\tau)$ in our notation) is shown to be intrinsic for both approaches: mathematical relations between informational states of H_∞ and set-membership state estimation are described in [5, Lemma 6.2.]. Comparisons of set-membership estimators with H_2/H_∞ filters for linear DAEs are presented in [19], provided $F_k \equiv E$. Let us consider connections to H_2-filters in details. In [10] the authors derive the Kalman's recursion to DAE from a deterministic least square fitting problem. Assuming $\operatorname{rank}\left[\begin{smallmatrix} F_k \\ H_k \end{smallmatrix}\right] \equiv n$, they prove that the optimal estimation

[7] $x \mapsto \langle Ax, x\rangle - 2\langle x, q\rangle + c$ is convex iff A is a symmetric non-negative matrix.

[8] Since $\operatorname{rank}\left[\begin{smallmatrix} F_k \\ H_k \end{smallmatrix}\right] < 2n$, it follows that results of [10] are not applicable.

$\hat{x}_{i|k}$ can be found from

$$\hat{x}_{k|k} = P_{k|k}F'_k A_{k-1}C_{k-1}\hat{x}_{k-1|k-1} + P_{k|k}H'_k R_k y_k,$$
$$\hat{x}_{0|0} = P_{0|0}H'_0 R_0 y_0, A^{-1}_k = S^{-1}_k + C_k P_{k|k} C'_k$$
$$P^{-1}_{k|k} = F'_k A_{k-1}F_k + H'_k R_k H_k, P^{-1}_{0|0} = F'_0 S F_0 + H'_0 R_0 H_0$$

Corollary 1 *Let* $r_0 = H'_0 R_0 y_0$. *If* $\operatorname{rank}\left[\begin{smallmatrix} F_k \\ H_k \end{smallmatrix}\right] \equiv n$ *then* $I_k = 0$ *and* $P^+_k r_k = \hat{x}_{k|k}$.

PROOF. Let us set $R_k = E, S = E, S_k = E$ for simplicity. The proof is by induction on k. For $k = 0$, $P_{0|0} = P^{-1}_0$. The induction hypothesis is $P_{k-1|k-1} = P^{-1}_{k-1}$. Suppose $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times n}$, $B = B' > 0$; then

$$A(A'A + B^{-1})^{-1} = (E + ABA')^{-1}AB \qquad (19)$$

Using (19) we get $ABA' = [E + ABA']A[A'A + B^{-1}]^{-1}A'$. Combining this with the induction assumption we get $E + C_{k-1}P_{k-1|k-1}C'_{k-1} = E + [E + C_{k-1}P_{k-1|k-1}C'_{k-1}]C_{k-1}[P_{k-1} + C'_{k-1}C_{k-1}]^{-1}C'_{k-1}$. By simple calculation it follows from the previous equality that $E - C_{k-1}(P_{k-1} + C'_{k-1}C_{k-1})^{-1}C'_{k-1} = (E + C_{k-1}P_{k-1|k-1}C'_{k-1})^{-1}$ Using this and definitions of P_k, $P_{k|k}$, we get $P^{-1}_k = P_{k|k}$.
$P^{-1}_0 r_0 = \hat{x}_{0|0}$ due to corollary assumption. Suppose that $P^{-1}_{k-1}r_{k-1} = \hat{x}_{k-1|k-1}$. The induction hypothesis, equality $P^{-1}_k = P_{k|k}$ and (19) imply $(E + C_{k-1}P_{k-1|k-1}C'_{k-1})^{-1}C_{k-1}\hat{x}_{k-1|k-1} = C_{k-1}(C'_{k-1}C_{k-1} + P_{k-1})^{-1}_{k-1}r_{k-1}$. Combining this with definitions of $\hat{x}_{k|k}$, r_k we obtain $\hat{x}_{k|k} = P^{-1}_k r_k$. This concludes the proof.

4 Conclusion

We describe a set-membership state estimation approach for a linear operator equation with uncertain disturbance restricted to belong to a convex bounded closed subset of abstract Hilbert space. It is based on the notion of an a posteriori set [16] $\mathscr{G}(y)$, informational set [5] and the notion of the minimax observable subspace for the pair (L, H). The latter is new for the set-membership state estimation framework. It leads to nontrivial new results in set-membership state estimation for linear non-causal DAEs: we present new equations describing the dynamics of the minimax recursive estimator for discrete-time non-causal DAEs. We prove that these equations are consistent with the main results already established for regular DAEs. We illustrate benefits of considering non-causality in the state equation, applying our approach to a linear filtration problem with unbounded noise.

Acknowledgements

It is a pleasure to thank Prof. A.Nakonechny and Dr. V.Pichkur for insightful discussions. I also thank to Drs. V.Mallet, J.A. Hosking and my anonymous referees for their help with improving the presentation of this paper.

References

[1] A. Albert. *Regression and the Moor-Penrose pseudoinverse.* Acad. press, N.Y., 1972.

[2] E. Angel and R. Bellman. *Dynamic programming and partial differential equations.* Acad. press N.Y., 1972.

[3] T. Başar and P. Bernhard. *H_∞-optimal Control and Related Minimax Design Problems.* Springer, 1995.

[4] A. Balakrishnan. *Applied functional analysis.* Springer, 1976.

[5] J. S. Baras and A.B. Kurzhanski. Nonlinear filtering: The set-membership and the H_∞ techniques. In *Proc. 3rd IFAC Symp.Nonlinear Control Sys.Design.* Pergamon, 1995.

[6] D. Bender and A. Laub. The linear-quadratic optimal regulator for descriptor system: discrete-time case. *Automatica*, 23:71–85, 1987.

[7] D.P. Bertsekas and I. B. Rhodes. Recursive state estimation with a set-membership description of the uncertainty. *IEEE Trans. Automat. Contr.*, AC-16:117–128, 1971.

[8] Z. Deng and Liu Y. Descriptor Kalman estimators. *Int. J. of System Science*, 30:1205–1212, 1999.

[9] I. Ekeland and R. Temam. *Convex analysis and variational problems.* North-Holland, Amsterdam, 1976.

[10] J.Y. Ishihara, M.H. Terra, and J.C.T. Campos. Optimal recursive estimation for discrete-time descriptor systems. *Int. J. of System Science*, 36(10):1–22, 2005.

[11] T. Kato. *Perturbation theory for linear operators.* Springer-Verlag, 1966.

[12] A. Kurzhanski and I. Valyi. *Ellipsoidal Calculus for Estimation and Control.* Birkhäuser, Boston, 1997.

[13] F. Lewis. A survey of linear singular systems. *Circuits Systems Signal Process*, 5(1), 1986.

[14] M. Milanese and R. Tempo. Optimal algorithms theory for robust estimation and prediction. *IEEE Trans. Autom. Contr.*, 30(8):730–738, Aug 1985.

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

[16] A. Nakonechny. Minimax estimation of functionals defined on solution sets of operator equations. *Arch.Math. 1, Scripta Fac. Sci. Nat. Ujer Brunensis*, 14:55–60, 1978.

[17] R. Nikoukhah, S.L. Campbell, and F. Delebecque. Kalman filtering for general discrete-time linear systems. *IEEE Transactions on Automatic Control*, 44:1829–1839, 1999.

[18] R. Rockafellar. *Convex analysis.* Princeton, 1970.

[19] A. Sayed. A framework for state-space estimation with uncertain models. *IEEE Trans. Autom. Contr.*, 46:998–1013, 2001.

[20] J.S. Shamma and Tu Kuang-Yang. Set-valued observers and optimal disturbance rejection. *IEEE Trans. Autom. Contr.*, 44(2):253–264, Feb 1999.

[21] R. Tempo. Robust estimation and filtering in the presence of bounded noise. *IEEE Trans. Autom. Contr.*, 33(9):864–867, 1988.

[22] S. Xu and J. Lam. Reduced-order H_∞ filtering for singular systems. *System & Control Letters*, 56(1):48–57, 2007.

[23] H. Zhang, L. Xie, and Y. Soh. Risk-sensitive filtering, prediction and smoothing for discrete-time singular systems. *Automatica*, 39:57–66, 2003.

[24] S. Zhuk. Closedness and normal solvability of an operator generated by a degenerate linear differential equation with variable coefficients. *Nonlin. Oscillations*, 10:1–18, 2007.

[25] S. Zhuk. State estimation for a dynamical system described by a linear equation with unknown parameters. *Ukr. Math. J.*, 61(2):178–194, 2009.