RECURSIVE STATE ESTIMATION FOR NONCAUSAL
DISCRETE-TIME DESCRIPTOR SYSTEMS UNDER
UNCERTAINTIES

SERHIY M. ZHUK

Abstract. This paper describes a method for the online state estimation of systems described by a general class of linear noncausal time-varying difference descriptor equations subject to uncertainties. The method is based on the notions of a linear minimax estimation and an index of causality introduced here for singular difference equations. The online minimax estimator is derived by the application of the dynamical programming and Moore’s pseudoinverse theory to the minimax estimation problem. It coincides with Kalman’s filter for regular systems. A numerical example of the state estimation for 2D noncasual descriptor system is presented.

Keywords Kalman filtering, online state observer, guaranteed estimation, descriptor systems, singular systems, DAEs.

1. Introduction

There is a number of physical and engineering objects most naturally modelled as systems of differential and algebraic equations (DAEs) or descriptor systems: microwave circuits [1], flexible-link planar parallel platforms [2] and image recognition problems (noncasual image modelling) [4]. DAEs arise in economics [5]. Also nonlinear differential-algebraic systems are studied with help of DAEs via linearization: a batch chemical reactor model [3].

On the other hand there are many papers devoted to the mathematical processing of data obtained from the measuring device during an experiment. In particular, a problem of the observer design for discrete-time descriptor systems was studied in the [5]-[8], the guaranteed state estimation for a linear dynamical systems was investigated in the [9]. In the [6] the authors derive a so-called ”3-block” form for the optimal filter and a corresponding 3-block Riccati equation using a maximum likelihood approach. A filter is obtained for a general class of time-varying descriptor models. The measurements are supposed to contain a noise with Gauss’es distribution. The obtained recursion is stated in terms of the 3-block matrix pseudoinverse.

In the [7] the filter recursion is represented in terms of a deterministic data fitting problem solution. The authors introduce an explicit form of the 3-block matrix pseudoinverse for a descriptor system with a special structure: so their filter coincides with obtained in the [6].

In this paper we study an observer design problem for a general class of linear noncasual time-varying descriptor models with no restrictions to a system structure. Suppose we are given an exact mathematical model of some real process and the vector $x_k$ describes the system output at the moment $k$ in the corresponding state space of the system. Also the successive measurements $y_0 \ldots y_k \ldots$ of the system
output $x_k$ are supposed to be available with the noise $g_0\ldots g_k\ldots$ of an uncertain nature\(^1\). Further assume that the system input $f_k$, start point $q$ and noise $g_k$ are arbitrary elements of the given set $G$. The aim of this paper is to design a minimax observer $k \mapsto \hat{x}_k$ that gives an online guaranteed estimation of the output $x_k$ on the basis of measurements $y_k$ and the structure of $G$. In [8] minimax estimations were derived from the 2-point boundary value problem with the conditions at $i = 0$ (start point) and $i = k$ (end point). Hence a recalculation of the whole history $\hat{x}_0\ldots\hat{x}_k$ is required if the moment $k$ changes. Here we derive the observer $(k, y_k) \mapsto \hat{x}_k$ by applying dynamical programming methods to the minimax estimation problem similar to posed in the [8]. We construct a map $\hat{x}$ that takes $(k, y_k)$ to $\hat{x}_k$ making it possible to assign a unique sequence of estimations $\hat{x}_0\ldots\hat{x}_k\ldots$ to the given sequence of observations $y_0\ldots y_k\ldots$ in the real time. A resulting filter recursion is stated in terms of the pseudoinverse of positive semi-defined $n \times n$-matrices.

2. Minimax estimation problem

Assume that $x_k \in \mathbb{R}^n$ is described by the equation

$$F_{k+1}x_{k+1} - C_kx_k = f_k, k = 0, 1, \ldots,$$

with the initial condition

$$F_0x_0 = q,$$

and $y_k$ is given by

$$y_k = H_kx_k + g_k, k = 0, 1, \ldots,$$

where $F_k, C_k$ are $m \times n$-matrices, $H_k$ is $p \times n$-matrix. Since we deal with a descriptor system we see that for any $k$ there is a set of vectors $x_1^{0}\ldots x_1^{0}$ satisfying (1) while $f_i = 0, q = 0$. Thus the undefined inner influence caused by $x_1^{0}\ldots x_1^{0}$ is possible to appear in the systems output. Also we suppose the initial condition $q$, input $\{f_k\}$ and noise $\{g_k\}$ to be unknown elements of the given set\(^2\)

$$G(q, \{f_k\}, \{g_k\}) = (Sq, q) + \sum_{0}^{\infty} (S_kf_k, f_k) + (R_kg_k, g_k) \leq 1$$

where $S, S_k, R_k$ are some symmetric positive-defined weight matrices with the appropriate dimensions. The trick is to fix any $N$-partial sum of (4) so that $(q, \{f_k\}, \{g_k\})$ belongs to

$$\mathcal{G}^N := \{(q, \{f_k\}, \{g_k\}) : (Sq, q) + \sum_{k=0}^{N-1} (S_kf_k, f_k) + \sum_{k=0}^{N} (R_kg_k, g_k) \leq 1\}$$

Then we derive the estimation $\hat{x}_N = v(N, y_N, \hat{x}_{N-1})$ considering a minimax estimation problem for $\mathcal{G}^N$. Lets denote by $\mathcal{N}$ a set of all $(\{x_k\}, q, \{f_k\})$ such that (1) is held. The set $\mathcal{G}^N_y$ is said to be a-\textit{posteriori set}, where

$$\mathcal{G}^N_y := \{\{x_k\} : (\{x_k\}, q, \{f_k\}) \in \mathcal{N}, (q, \{f_k\}, \{y_k - H_kx_k\}) \in \mathcal{G}^N\}$$

---

\(^1\)For instance we do not have a-priory information about its distribution.

\(^2\)Here and after $(\cdot, \cdot)$ denotes an inner product in an appropriate euclidean space, $\|x\| = (x, x)^{\frac{1}{2}}$. 
It follows from the definition that \( \mathcal{G}_y^N \) consists of all possible \( \{x_k\} \) causing an appearance of given \( \{y_k\} \) while \( q, \{f_k\}, \{g_k\} \) runs through \( \mathcal{G}^N \). Thus, it’s naturally to look for \( x_N \) estimation only among the elements of \( P_N(\mathcal{G}_y^N) \), where \( P_N \) denotes the projection that takes \( \{x_0 \ldots x_N\} \) to \( x_N \).

**Definition 1.** A linear function \( \widehat{(\ell, x_N)} \) is called a minimax a-posteriori estimation if the following condition holds:

\[
\inf_{\{x_k\} \in \mathcal{G}_y^N} \sup_{\{x_k\} \in \mathcal{G}_y^N} |(\ell, x_N) - (\ell, \widehat{x_N})| = \sup_{\{x_k\} \in \mathcal{G}_y^N} |(\ell, x_N) - (\ell, \widehat{x_N})|
\]

The non-negative number

\[
\hat{\sigma}(\ell, N) = \sup_{\{x_k\} \in \mathcal{G}_y^N} |(\ell, x_N) - (\ell, \widehat{x_N})|
\]

is called a minimax a-posteriori error in the direction \( \ell \). A map

\[N \mapsto I_N = \dim(\ell \in \mathbb{R}^n : \hat{\sigma}(\ell, N) < +\infty)\]

is called an index of causality for the pair of systems (1)-(3).

Denote by \( k \mapsto Q_k \) a recursive map that takes each \( k \in \mathbb{N} \) to the matrix \( Q_k \), where

\[
\begin{align*}
Q_k &= H_k^T R_k H_k + F_k^T [S_{k-1} - S_{k-1} C_{k-1} W_{k-1}^+ C_{k-1}^T S_{k-1}] F_k, \\
Q_0 &= F_0^T S F_0 + H_0^T R_0 H_0, W_k = Q_k + C_k^T S_k C_k
\end{align*}
\]

Let \( k \mapsto r_k \) be a recursive map that takes each natural number \( k \) to the vector \( r_k \in \mathbb{R}^n \), where

\[
\begin{align*}
r_k &= F_k^T S_{k-1} C_{k-1} W_{k-1}^+ r_{k-1} + H_k^T r_k y_k, \\
r_0 &= H_0^T r_0 y_0
\end{align*}
\]

and to each number \( i \in \mathbb{N} \) assign the number \( \alpha_i \), where

\[
\begin{align*}
\alpha_i &= \alpha_{i-1} + (R_i y_i, y_i) - (W_{i-1}^+ r_{i-1}, r_{i-1}), \\
\alpha_0 &= (S g, g) + (R_0 y_0, y_0)
\end{align*}
\]

The main result of this paper is formulated in the next theorem.

**Theorem 1** (minimax recursive estimation). Suppose we are given a natural number \( N \) and a vector \( \ell \in \mathbb{R}^n \). Then a necessary and sufficient condition for a minimax a-posteriori error \( \hat{\sigma}(\ell, N) \) to be finite is that

\[
Q_N^+ Q_N \ell = \ell
\]

Under this condition we have

\[
\hat{\sigma}(\ell, N) = [1 - \alpha_N + (Q_N^+ r_N, r_N)]^{\frac{1}{2}} (Q_N^+ \ell, \ell)^{\frac{1}{2}}
\]

and

\[
(\ell, x_N) = (\ell, Q_N^+ r_N)
\]

**Corollary 1.** The index of causality \( I_N \) for the pair of systems (1)-(3) can be represented as \( I_N = \text{rank}(Q_N) \).
Corollary 2 (minimax observer). The online minimax observer is given by \( k \mapsto \hat{x}_k = Q_k^{-1}r_k \) and \(^3\)

\[
\hat{\rho}(N) = \min_{\{x_k\} \in \mathcal{G}_N} \max_{(\hat{x}_k) \in \mathcal{G}_N} \|x_N - \hat{x}_N\|^2 = \\
\frac{1 - \alpha_N + (Q_N\hat{x}_N, \hat{x}_N)}{\min_i \{\lambda_i(N)\}}
\]

(13)

where \( \lambda_i(N) \) are eigen values of \( Q_N \). In this case all possible realisations of (1) state vector \( x_N \) fill the ellipsoid \( P_N(\mathcal{G}_N^y) \subset \mathbb{R}^n \), where

\[
P_N(\mathcal{G}_N^y) = \{ x : (Q_Nx, x) - 2(Q_N\hat{x}_N, x) + \alpha_N \leq 1 \}
\]

Remark 1. If \( \lambda_{\min}(H_k^*R_kH_k) \) grows for \( k = i, i+1, \ldots \) then the minimax estimation error \( \hat{\rho}(k) \) becomes smaller causing \( \hat{x}_k \) to get closer to the real state vector \( x_k \).

In [7] Kalman’s filtering problem for descriptor systems was investigated from the deterministic point of view. Authors recover Kalman’s recursion to the time-variant descriptor system by a deterministic least square fitting problem over the entire trajectory: find a sequence \( \{\hat{x}_{0|k}, \ldots, \hat{x}_{k|k}\} \) that minimises the following fitting error cost

\[
J_k(\{x_{i|k}\}_0^k) = \|F_0x_{0|k} - g\|^2 + \|y_0 - H_0x_{0|0}\|^2 + \\
\sum_{i=1}^k \|F_ix_{i|k} - C_{i-1}x_{i-1|k}\|^2 + \|y_i - H_ix_{i|k}\|^2
\]

assuming that the rank \( F_k = n \). According to [7, p.8] the successive optimal estimates \( \{\hat{x}_{0|k}, \ldots, \hat{x}_{k|k}\} \) resulting from the minimisation of \( J_k \) can be found from the recursive algorithm

\[
\hat{x}_{k|k} = F_kx_{k|k} + P_{k|k}H_ky_k, \hat{x}_{0|0} = P_{0|0}(F_0^tg + H_0^ty_0),
\]

\[
P_{k|k} = (F_k^t(E + C_{k-1}P_{k-1|k-1}C_{k-1}^t))^{-1}F_k + H_k^tH_k)^{-1},
\]

\[
P_{0|0} = (F_0^tg + H_0^ty_0)^{-1}
\]

Corollary 3 (Kalman’s filter recursion). Suppose the rank \( F_k = n \), and let \( k \mapsto r_k \) be a recursive map that takes each natural number \( k \) to the vector \( r_k \in \mathbb{R}^n \), where

\[
r_k = H_k^*y_k + F_k^tC_k^{-1}(C_{k-1}^tC_{k-1} + Q_{k-1})^{-1}r_{k-1},
\]

\[
r_0 = F_0^tg + H_0^ty_0
\]

Then \( Q_k^{-1}r_k = \hat{x}_{k|k} \) for each \( k \in \mathbb{N} \), where \( \hat{x}_{k|k} \) is given by (15) and \( I_k = n \).

3. Example

Let us set \( H_0 = \begin{bmatrix} \frac{1}{1000} & 0 & 0 \\ 1 & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix} \), \( F_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \),

\[
C_k = \begin{bmatrix} \frac{1}{100} & \frac{1}{100} & 0 \\ \frac{1}{100} & \frac{1}{100} & 0 \\ 0 & 0.005 & 0 \end{bmatrix},
\]

\[
H_k \equiv \begin{bmatrix} k + \frac{1}{100} & 0 & 0 \\ 0 & \frac{1}{100} & 0 \end{bmatrix}
\]

\[
\frac{1}{100} \frac{1}{100} 0 \frac{1}{100} \frac{1}{100} 0 \frac{1}{100} \frac{1}{100} 0 \frac{1}{100} \frac{1}{100} 0 \frac{1}{100} \frac{1}{100} 0 \frac{1}{100} \frac{1}{100} 0 \frac{1}{100} \frac{1}{100}
\]

\[3\text{We assume here that } \frac{1}{0} = +\infty.\]
where \( q(k) = 1 \) if \( k \) is odd and otherwise \( q(k) = 0 \). We derive the output \( x_k \) of (1) and \( y_k \) assuming \( f_k, g_k \) to be bounded vector-functions on the whole real axis. Also we set
\[
R_k = \text{diag}\{\frac{1}{11(k+1)}, \frac{1}{22(k+1)}, \frac{1}{33(k+1)}, \frac{1}{44(k+1)}\},
\]
\[
S_k = \text{diag}\{\frac{1}{35(k+1)}, \frac{1}{70(k+1)}\},
\]
\[
S = \text{diag}\{\frac{1}{60}, \frac{1}{120}\}.\]

We derive \( \hat{x}_k \) from (8) and \( \hat{\sigma}(e_i, k) \) from (11), \( e_i \) - i-ort. Note that the rank \( F_{2k+1} < 3 \) and \( I_{2k} = 3, I_{2k+1} < 3 \). Thus \( \hat{x}_{3,2k+1} = 0 \),
\[
[1 - \alpha_{2k+1} + (Q_{2k+1}^+ r_{2k+1}, r_{2k+1})] \hat{\sigma}(Q_{2k+1}^+, \ell) \hat{x}_{2k+1} = 0
\]
but \( |x_{3,2k+1} - \hat{x}_{3,2k+1}| > 0 \). The dynamics of \( x_{i,k}, \hat{x}_{i,k}, |x_{i,k} - \hat{x}_{i,k}| \) and \( \hat{\sigma}(e_i, k) \) is described by figures 1-2.

![Figure 1](image_url)

**Figure 1.** \( N = 100 \), output \( x_{i,k} \) (solid) and observer \( \hat{x}_{i,k} \) (dashed) to the left;
Figure 2. $N = 100$, real estimation error $|x_{i,k} - \hat{x}_{i,k}|$ (dashed) and minimax error $\hat{\sigma}(e_i, k)$ (solid) to the right.

Acknowledgements It is a pleasure to thank Prof. A.Nakonechniy and Dr. V.Pichkur for insightful discussions about the key ideas presented in this paper.

Appendix A. Proofs.

Proof. Proof of Theorem 1. By definition, put

$$H = \begin{pmatrix} H_0 & 0_{p_0} & \ldots & 0_{p_n} \\ 0_{p_0} & H_1 & \ldots & 0_{p_n} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{p_0} & 0_{p_0} & \ldots & H_N \end{pmatrix}, F = \begin{pmatrix} F_0 & 0_{m_0} & 0_{m_0} & \ldots & 0_{m_n} & 0_{m_n} \\ -C_0 & F_1 & 0_{m_0} & \ldots & 0_{m_n} & 0_{m_n} \\ 0_{m_0} & -C_1 & F_2 & \ldots & 0_{m_n} & 0_{m_n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0_{m_0} & 0_{m_0} & \ldots & -C_{N-1} & F_N \end{pmatrix}$$
$X = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix}, \ Y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix},$

$\mathcal{F} = \begin{bmatrix} q \\ f_0 \\ f_1 \\ \vdots \\ f_{N-1} \end{bmatrix}, \ G = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_N \end{bmatrix}.$

By direct calculation we obtain $(\ell, x_N) = (\mathcal{L}, X),$

$$\mathcal{G}_0^N = \{X : \|F_X\|^2 + \|Y - H_X\|^2 \leq 1\},$$

where $\|F\|^2 = (Sq, q) + \sum_{0}^{N-1}(S_k f_k, f_k),$ $\| \cdot \|_2$ is induced by $R_k$ on the same way. This implies

$$\sup_{(x_k) \in \mathcal{G}_0^N} |(\ell, x_N - \tilde{x}_N)| = \sup_{X \in \mathcal{G}_0^N} |(\mathcal{L}, X) - (\mathcal{L}, \tilde{X})|$$

Denote by $L$ the set $R[\gamma' \ H' \ W']$. We obviously get

$$\mathcal{L} \in L \iff \sup_{X \in \mathcal{G}_0^N} |(\mathcal{L}, X) - (\mathcal{L}, \tilde{X})| < +\infty$$

The application of Corollary 4 yields (10). Consider a vector $\mathcal{L} \in L$. Clearly

$$\inf_{X \in \mathcal{G}_0^N} (\mathcal{L}, X) \leq (\mathcal{L}, X) \leq \sup_{X \in \mathcal{G}_0^N} (\mathcal{L}, X), X \in \mathcal{G}_0^N$$

Let $c$ denotes $\frac{1}{2}(\sup_{X \in \mathcal{G}_0^N} (\mathcal{L}, X) + \inf_{X \in \mathcal{G}_0^N} (\mathcal{L}, X))$. Therefore

$$\sup_{X \in \mathcal{G}_0^N} |(\mathcal{L}, X) - (\mathcal{L}, \tilde{X})| =$$

$$\frac{1}{2} (s(\mathcal{L}|\mathcal{G}_0^N) + (s(-\mathcal{L}|\mathcal{G}_0^N)) + |c - (\mathcal{L}, \tilde{X})|$$

hence

$$\hat{\sigma}(\ell, N) = \frac{1}{2} (s(\mathcal{L}|\mathcal{G}_0^N) + s(-\mathcal{L}|\mathcal{G}_0^N)), \tag{17}$$

$$\hat{\sigma}(\ell, x_N) = \frac{1}{2} (s(\mathcal{L}|\mathcal{G}_0^N) - s(-\mathcal{L}|\mathcal{G}_0^N)),$$
Le ma 2.

(20) \[ s(\mathcal{L}|g_0^N) = \begin{cases} \sqrt{\beta_N(N^+Q_N \ell, \ell)^2} [E - Q_N^+Q_N] \ell = 0, \\ +\infty, [E - Q_N^+Q_N] \ell \neq 0 \end{cases} \]

\[ \square \]

Let \( r_k \) denote \( \mathbb{R}^n \)-valued recursive map

\[ r_k = F_k^s(S_{k-1} - S_{k-1}C_{k-1}P_{k-1}^+C_{k-1}S_{k-1})f_{k-1} + \]

(21) \[ F_k^sS_{k-1}C_{k-1}W_{k-1}^r r_{k-1} + H_k^rR_k y_k, \]

\[ r_0 = F_0^sS_q + H_0^rR_0 y_0, \]

and set

\[ J(\{x_k\}) = \|F_0 x_0 - g\|^2_S + \|y_0 - H_0 x_0\|^2_0 \]

(22) \[ \sum_{k=1}^N \|F_k x_k - C_{k-1} x_{k-1} - f_{k-1}\|^2_{k-1} + \|y_k - H_k x_k\|^2_k \]

where \( \|g\|^2_S = (S g, g), \|f_k\|^2_k = (S_k f_k, f_k), \|y_k\|^2_i = (R_i y_i, y_i). \]

Le ma 3. Let \( x \mapsto \hat{x}_k \) be a recursive map that takes any \( k \in \mathbb{N} \) to \( \hat{x}_k \in \mathbb{R}^n \), where

\[ \hat{x}_k = F_k^s(C_S^k(F_{k+1}\hat{x}_{k+1} - f_k) + r_k), \]

Then

\[ \min_{\{x_k\}} J(\{x_k\}) = J(\{\hat{x}_k\}) \]

\[ \Phi(x_0) := \|F_0 x_0 - g\|^2_S + \|y_0 - H_0 x_0\|^2_0 \]

\[ \Phi_i(x_i, x_{i+1}) := \|F_{i+1} x_{i+1} - C_i x_i - f_i\|^2_i + \|y_{i+1} - H_{i+1} x_{i+1}\|^2_{i+1} \]

Then we obviously get

(23) \[ J(\{x_k\}) = \Phi(x_0) + \sum_{i=0}^{N-1} \Phi_i(x_i, x_{i+1}) \]

Let us apply a modification of Bellman’s method\(^4\) to the nonlinear programming task

\[ J(\{x_k\}) \to \min_{\{x_k\}} \]

By definition put

\[ \ell_1(x_1) := \min_{x_0} \{\Phi(x_0) + \Phi_0(x_0, x_1)\} \]

Using (7) and (21) one can get

\[ \Phi(x_0) = \langle Q_0 x_0, x_0 \rangle - 2(r_0, x_0) + \alpha_0 \geq 0, \]

where \( \alpha_0 := \|g\|^2_S + \|y_0\|^2_0 \)

On the other hand it’s clear that

\[ \ell_1(x_1) = \Phi(\hat{x}_0) + \Phi_0(\hat{x}_0, x_1) = \langle Q_1 x_1, x_1 \rangle - 2(r_1, x_1) + \alpha_1 \geq 0, \]

\[ \alpha_1 := \alpha_0 + \|y_1\|^2_1 + \|f_0\|^2_0 - (P_0^+(r_0 - C_0^0 S_0 f_0), r_0 - C_0^0 S_0 f_0) \]

\[^4\text{So-called “Kievskiy venyk” method}\]
Considering \( \ell_1(x_1) \) as an induction base and assuming that
\[
\ell_{i-1}(x_{i-1}) = \min_{x_{i-2}} \{ \Phi_{i-2}(x_{i-2}, x_{i-1}) + \ell_{i-2}(x_{i-2}) \} = (Q_{i-1}x_{i-1}, x_{i-1}) - 2(r_{i-1}, x_{i-1}) + \alpha_{i-1}
\]
now we are going to prove that
\[
\ell_i(x_i) = \min_{x_{i-1}} \{ \Phi_{i-1}(x_{i-1}, x_i) + \ell_{i-1}(x_{i-1}) \} = (Q_i x_i, x_i) - 2(r_i, x_i) + \alpha_i
\]
(24)

Note that [10] for any convex function \((x, y) \mapsto f(x, y)\)
\[
y \mapsto \min \{ f(x, y) : P(x, y) = y \}, P(a, b) = b
\]
is convex. Thus taking into account the definition of \( \ell_1(x_1) \) one can prove by induction that \( \ell_{i-1} \) is convex and
\[
\Phi_{i-1}(x_{i-1}, x_i) + \ell_{i-1}(x_{i-1}) \geq 0
\]
Hence \( Q_{i-1} \geq 0 \), the set of global minimums \( \Psi_{i-1} \) of the quadratic function
\[
x_{i-1} \mapsto \Phi_{i-1}(x_{i-1}, x_i) + (Q_{i-1}x_{i-1}, x_{i-1}) - 2(r_{i-1}, x_{i-1}) + \alpha_{i-1}
\]
is non-empty and \( \hat{x}_{i-1} \in \Psi_{i-1} \), where
\[
\hat{x}_{i-1} = (Q_{i-1} + C_{i-1}S_{i-1}C_{i-1})^+(C_{i-1}S_{i-1}(F_i x_i - f_i + r_{i-1}))
\]
This implies
\[
\ell_i(x_i) = \Phi_{i-1}(\hat{x}_{i-1}, x_i) + \ell_{i-1}(\hat{x}_{i-1}) = (Q_i x_i, x_i) - 2(r_i, x_i) + \alpha_i,
\]
where
\[
\alpha_i = \alpha_{i-1} + (R_i y_i, y_i) + (S_{i-1}f_{i-1}, f_{i-1}) - (P_{i-1}^+(r_{i-1} - C_{i-1}S_{i-1}f_{i-1}), r_{i-1} - C_{i-1}S_{i-1}f_{i-1}).
\]
Therefore, we obtain
\[
\min_{x_N} \ell_N(x_N) = \ell_N(\hat{x}_N) = \alpha_N - (r_N, Q_N^N r_N), \hat{x}_N = Q_N^N r_N
\]
so that \( \min_{x_N} \) \( J(\{x_k\}) = J(\{\hat{x}_k\}) \).

**Corollary 4.** Suppose \( \mathcal{L} = [0 \ldots \ell] ; \) then
\[
\mathcal{L} \in \mathcal{R}[v' \ w'] \iff [E - Q_N^N Q_N] \ell = 0
\]
and
\[
\|[v' \ w']^+ \mathcal{L}\|^2 = (Q_N^N \ell, \ell)
\]

**Proof.** Suppose \( S_k = E, R_k = E \) for a simplicity. If \( \mathcal{L} \in \mathcal{R}[v' \ w'] \) then
\[
F_N' z_N + H_N' u_N = \ell, \quad F_k' z_k + H_k' u_k - C_k' \hat{x}_{k+1} = 0 \quad (\ast)
\]
for some \( z_k \in \mathbb{R}^m, u_k \in \mathbb{R}^p \). Let’s find the projection \( \{(\hat{z}_k, \hat{u}_k)\}_{k=0}^N \) of the vector \( \{(z_k, u_k)\}_{k=0}^N \) onto the range of the matrix \( [F_H] \). Lemma 3 implies
\[
\hat{z}_0 = F_0 \hat{x}_0, \hat{z}_k = F_k \hat{x}_k - C_{k-1} \hat{x}_{k-1}, \hat{u}_k = H_k \hat{x}_k, \quad (\ast\ast)
\]
\(^5\)The function \( x \mapsto (Ax, x) - 2(x, q) + c \) is convex iff \( A = A' \geq 0 \).
\(^6\)The vector \( \hat{x}_{i-1} \) has the smallest norm among other points of the minimum.
where
\[
\hat{x}_k = P_k^e (C_k^t F_{k+1} \hat{x}_{k+1} + r_k - C_k^t z_{k+1}), \hat{x}_N = Q_N^e r_N,
\]
\[
r_k = F_k^e C_{k-1} P_{k-1}^+ r_{k-1} + F_k^e (E - C_{k-1} P_{k-1}^+ C_k^t) z_k + H_k^e u_k, r_0 = F_0^e z_0 + H_0^e u_0, P_k = C_k^t C_k + Q_k
\]

(*) implies
\[
r_k = C_k^t z_{k+1}, k = 0, \ldots, N - 1, r_N = \ell \text{ thus } \hat{x}_N = Q_N^e \ell, \hat{x}_k = P_k^e C_k^t F_{k+1} \hat{x}_{k+1} \text{ or } \hat{x}_k = \Phi(k, N) Q_N^e \ell, \Phi(k, N) = P_k^e C_k^t F_{k+1} \Phi(k+1, N), \Phi(s, s) = E
\]
Combining this with (**) we obtain
\[
\hat{z}_k = (F_k \Phi(k, N) - C_{k-1} \Phi(k-1, N)) Q_N^e \ell, \hat{u}_k = H_k \Phi(k, N) Q_N^e \ell, \hat{z}_0 = F_0 \Phi(0, N) Q_N^e \ell
\]
By definition, put \( U(0) = Q_0, \)
\[
U(k) = \Phi'(k-1, k) U(k-1) \Phi(k-1, k) + H_k^e H_k + F_k (E - C_{k-1} P_{k-1}^+ C_k^t)^2 F_k
\]
It now follows that
\[
\| [v', w']^+ L \|^2 = \sum_{0}^{N} \| \hat{z}_N \|^2 + \| \hat{u}_N \|^2 = (U(N) Q_N^e \ell, Q_N^e \ell)
\]
It’s easy to prove by induction that \( Q_k = U(k). \)

Since
\[
L \in R[v', w']
\]
we obtain by substituting \( \hat{z}_k, \hat{u}_k \) into (*)
\[
F_N^e \hat{z}_N + H_N^e \hat{u}_N = \ell
\]
On the other hand (7) and (25) imply
\[
F_N^e \hat{z}_N + H_N^e \hat{u}_N = \ell \Rightarrow [E - Q_N^e Q_N] \ell = 0
\]
Suppose that \( [E - Q_N^e Q_N] \ell = 0. \) To conclude the proof we have to show that
\[
(\ell, x_N) = (Q_N^e \ell, Q_N x_N) = 0, \forall [x_0 \ldots x_N] \in \mathcal{N}[v, w]
\]
By induction, fix \( N = 0. \) If \( F_0 x_0 = 0, H_0 x_0 = 0, \text{ then } Q_0 x_0 = 0. \) We say that
\[
[x_0 \ldots x_k] \in \mathcal{N}[v, w] \text{ if }
\]
\[
F_0 x_0 = 0, H_0 x_0 = 0, F_s x_s = C_{s-1} x_{s-1}, H_s x_s = 0,
\]
Suppose \( Q_{k-1} x_{k-1} = 0, \forall [x_0 \ldots x_{k-1}] \in \mathcal{N}[v, w] \text{ and fix any } [x_0 \ldots x_k] \in \mathcal{N}[v, w]. \)
Then \( F_k x_k = C_{k-1} x_k, H_k x_k = 0. \) Combining this with (7) we obtain
\[
Q_k x_k = F_k^e (E - C_{k-1} P_{k-1}^+ C_{k-1}) C_{k-1} x_{k-1} \quad (*)
\]
We show that \( Q_k \geq 0 \) in the proof of Theorem 1. One can see that
\[
\begin{bmatrix}
C_{k-1} \\
Q_{k-1}^2
\end{bmatrix}^+ = \begin{bmatrix}
(C_{k-1}^e C_{k-1} + Q_{k-1})^e C_{k-1}^e C_{k-1} + Q_{k-1}^e Q_{k-1}^2
\end{bmatrix}
\]
Since
\[
\begin{bmatrix}
C_{k-1} \\
Q_{k-1}^2
\end{bmatrix} \begin{bmatrix}
C_{k-1} \\
Q_{k-1}^2
\end{bmatrix}^+ \begin{bmatrix}
C_{k-1} \\
Q_{k-1}^2
\end{bmatrix} x_{k-1} = \begin{bmatrix}
C_{k-1} \\
Q_{k-1}^2
\end{bmatrix} x_{k-1}
\]
Using this and (7),(15) we obviously get nothing to prove. The induction hypothesis is

\[ P_n \]

Proof. By simple calculation from the previous equality follows (26)

Proof of Corollary 3. The proof is by induction on this rule and then apply Corollary 4.

Combining (26) with the induction assumption we get the following

\[ A(S^{-1} + A' A)^{-1} = (E + ASA')^{-1} AS \]

Using (26) we get

\[ ASA' = [E + ASA'] A[A' A + S^{-1}]^{-1} A' \]

Combining (27) with the induction assumption we get the following

\[ E + C_{k-1} P_{k-1|k-1} C'_{k-1} = E + [E + C_{k-1} P_{k-1|k-1} C'_{k-1}] \times \]

\[ \times C_{k-1} [Q_{k-1} + C'_{k-1} C_{k-1}]^{-1} C'_{k-1} \]

By simple calculation from the previous equality follows

\[ E - C_{k-1} (Q_{k-1} + C'_{k-1} C_{k-1})^{-1} C'_{k-1} = (E + C_{k-1} P_{k-1|k-1} C'_{k-1})^{-1} \]

Using this and (7),(15) we obviously get \( Q_{k-1}^{-1} = P_{k|k} \).

It follows from the definitions that \( Q_{0|0}^{-1} = \tilde{x}_{0|0} \). Suppose that \( Q_{k-1}^{-1} r_{k-1} = \tilde{x}_{k-1|k-1} \). The induction hypothesis and (26) imply

\[ (E + C_{k-1} P_{k-1|k-1} C'_{k-1})^{-1} C_{k-1} \tilde{x}_{k-1|k-1} = C_{k-1} (C_{k-1} + Q_{k-1})^{-1} r_{k-1} \]
Combining this with (15), (16) and using \( Q_k^{-1} = P_{k|k} \) we obtain
\[
\hat{x}_{k|k} = Q_k^{-1}(F'_k C_{k-1}(C'_{k-1} C_{k-1} + Q_{k-1})^{-1} r_{k-1} + H'_k y_k)
\]
This concludes the proof.

**Proof.** Proof of Corollary 2. If \( I_k < n \) then \( \operatorname{rank}(Q) < n \) hence \( \lambda_{\min}(Q_k) = 0 \). In this case there is a direction \( \ell \in \mathbb{R}^n \) such that \( \hat{\sigma}(\ell, k) = +\infty \). So \( \hat{\rho}(k) = +\infty \). If \( I_k = n \) then we clearly have
\[
\min_{\{x_k\} \in \mathbb{G}^N} \max_{\{\hat{x}_k\} \in \mathbb{G}^N} \|x_N - \hat{x}_N\|^2 = 0
\]
\[
\min_{\|\ell\|=1} \max_{\{x_k\} \in \mathbb{G}^N} \{ \max_{\{\hat{x}_k\} \in \mathbb{G}^N} |(\ell, x_N - \hat{x}_N)| \}^2 = 0
\]
\[
\max_{\|\ell\|=1} \min_{\{x_k\} \in \mathbb{G}^N} \max_{\{\hat{x}_k\} \in \mathbb{G}^N} |(\ell, x_N - \hat{x}_N)|^2 = 0
\]
\[
[1 - \alpha_N + (Q^+_N r_N, r_N)] \max_{\|\ell\|=1} (Q^+_N \ell, \ell) = 0
\]
On the other hand Theorem 1 implies
\[
\max_{\{\hat{x}_k\} \in \mathbb{G}^N} \|\hat{x}_N - \hat{x}_N\|^2 = 0
\]
\[
\{ \max_{\|\ell\|=1} \max_{\{x_k\} \in \mathbb{G}^N} |(\ell, x_N - \hat{x}_N)| \}^2 = 0
\]
\[
\{ \max_{\|\ell\|=1} [1 - \alpha_N + (Q^+_N r_N, r_N)] \}^2 (Q^+_N \ell, \ell) = 0
\]
It follows now from \( I_N = n \) that \( \mathbb{G}^N_y \) is a bounded set.

The equality \( I_N = n \) implies \( [E - Q^+_N Q_N] = 0 \) for a given \( N \). It follows from Lemmas 1,2 that
\[
s(\ell | P_N(\mathbb{G}^N_y)) = s(P'_N \ell | \mathbb{G}^N_y) = s(L | \mathbb{G}^N_y) = (\ell, Q^+_N r_N) + \sqrt{\beta_N(\ell, \ell)\frac{1}{2}}
\]
for any \( \ell \in \mathbb{R}^n \). By Young’s theorem [10], (28), so that
\[
P_N(\mathbb{G}^N_y) = \{ x \in \mathbb{R}^n : (x, \ell) \leq s(\ell | P_N(\mathbb{G}^N_y)), \forall \ell \in \mathbb{R}^n \} = \{ x \in \mathbb{R}^n : \sup_{\ell} \{ (x, \ell) - (\ell, \hat{x}_N) - \sqrt{\beta_N(\ell, \ell)\frac{1}{2}} \} \leq 0 \} = \{ x \in \mathbb{R}^n : (Q_N x, x) - 2(Q_N \hat{x}_N, x) + \alpha_N \leq 1 \}
\]

**REFERENCES**

[1] *Favini A.*, *Vlasenko L.* On solvability of degenerate nonstationary differential-difference equations in Banach spaces.– Diff. & Int. Equn., 2001, V14, N7, p.883-896

[2] *James K. Mills* Dynamic modelling of a a flexible-link planar parallel platform using a substructuring approach.– Mechanism and Machine Theory 41 (2006) p.671-687

[3] *V.M. Becerra, P.D. Roberts, G.W. Griffiths* Applying the extended Kalman filter to systems described by nonlinear differential-algebraic equations.– Control Engineering Practice 9 (2001) p.267-281
[4] Hasan M.A. Noncausal image modelling using descriptor approach // IEEE Transactions on Circuits and Systems II. – 1995. – Vol. 2, 42. – P. 536-540.

[5] Dai L. Singular control systems.”— Lec. Notes in Control and Inform. Sciences, 1989, Berlin: Springer

[6] R. Nikoukhah, S. L. Campbell, and F. Delebecque Kalman Filtering for General Discrete-Time Linear Systems, IEEE Transactions on Automatic Control, 44, pp. 1829-1839, 1999.

[7] Ishihara J.Y., Terra M.H., Campos J.C.T. Optimal recursive estimation for discrete-time descriptor systems // Int. J. of System Science. – 2005. – Vol. 36, N10. – P. 1-22.

[8] S.Zhuk Minimax estimations for linear descriptor difference equations // www.arxiv.org

[9] Kyrychenko M., Nakonechny A. Minimax approach to recursive estimation of the linear dynamical systems // Cybernetics 1977 N5

[10] R. Rockafellar Convex analysis // Princeton University Press, 1970

Taras Shevchenko Kyiv National University, Faculty of Cybernetics, System Analysis Laboratory, Ukraine

E-mail address: Serhiy.Zhuk@gmail.com