UDC [681.5.015.4::519.85]:681.51

V. Gubarev, M. Mishchenko

INTERVAL STATE ESTIMATOR FOR LINEAR SYSTEMS WITH KNOWN STRUCTURE

Vyacheslav Gubarev

Space Research Institute under NAS of Ukraine and State Space Agency of Ukraine, Kyiv,
orcid: 0000-0001-6284-1866
v.f.gubarev@gmail.com

Mykhailo Mishchenko

National Technical University of Ukraine «Igor Sikorsky Kyiv Politechnic Institute»,
Institute for Applied Systems Analysis, Kyiv,
orcid: 0000-0001-6135-2569
mdmisch@firemail.cc, mdmisch@protonmail.com

It is often required to control a system whose state is not observable directly. Instead, there are indirect incomplete and noised measurements of its state. In such situation it is required to estimate current system's state from these indirect measurements first in order to control the system. For this purpose the Kalman filter is the long established and classical approach on estimation of linear system's state from indirect measurements. It is recursive by desin, and thus indirectly takes into account the whole previous history of measurements. Here we explore an alternative approach: estimation with measurements on a limited historic horizon. The article first discusses application of the generalized linear least squares (GLLS) estimator to this problem and conditions under which it is appropriate to use this method. For situations when it is not fully appropriate, we propose a way to represent the GLLS estimator as a quadratic cone programming problem which helps producing its modifications tuned for various nonstandard linear system designs. The article also explores various properties and behavior of the GLLS estimator and its modifications. For instance, it is completely expectable that such estimators demonstrate diferent precision with different number of historic measurements considered. Thus, application of the absolute condition number of the GLLS estimator to choosing an optimal horizon length was explored. It was demonstrated how the absolute condition number of GLLS, while being a hard limit on estimation precision, also limits expected value of error norm. Choice of the best horizon length was discussed from both of these points of view. For situations when best possible estimation precision is still not enough, a regularization method was proposed. Pros and cons of this regularization method and a way to make an informed choice regarding the degree of regularization was explored. The theoretical results were confirmed with computational experiments.

Keywords: linear system, state estimation, limited measurement historic horizon, linear least squares, quadratic cone programming.

Introduction

For state-space model-based feedback control design built with existing methods the knowledge of full state vector at each current time is required. In practice, this knowledge is often not available through direct measurements and the so-called output vector can be measured instead. This output vector depends on the actual current state through a linear equation and its dimension is less then state vector's. This means that control system should include a state estimator which takes output vector measurements and motion equation as its input. For deterministic case the Luenberger observer was proposed, which for precise model and data can recover insufficient information [1]. However, usage of the Kalman filter is currently more realistic and widespread approach in practical applications, as for process impacted by white noise gives the optimal state estimation. This method also obtained wide application in many other fields and not only in control design. As it applies to linear time-invariant discrete systems, the Kalman filter description is given in monographs [2]. Other approaches to state estimation are also considered along with the Kalman filter [3]. One of them is a guaranteed ellipsoidal state estimation, see for example [4–7]. Its distinctive feature is that instead of interpretation of stochastic noise and disturbances it uses nonstohastic approach where errors in data are represented as possible values' set. Comparatively not long ago it established a new trend in control: the so called "Model Predictive Control" (MPC). Number of its practical applications grows fast. [8] This approach is also model-based and equations of dynamic system are used to synthesize producing an optimal control for some future horizon. During the application of the synthesized control signals system's actual state is compared to the predicted one and further controls are corrected accordingly. If system's state is not measured directly, then it implies that some kind of state estimator is also required. Having dimensionality of output vector smaller than state vector's, the only way to make an estimation is to directly or indirectly consider multiple historical output vector's measurements. The Kalman filter's approach on this is to store essential data extracted from previous measurements in its internal variables. Another way to do it is to consider explicitly a specific number of prevous output vector measurements in an estimation, which is called the moving horizon estimation (MHE). The MHE problem has been studied by many authors, see [9–12].

One possible way of implementing such an estimator is presented in this paper. Features of estimation process and conditions of its applicability are considered.

1. Estimation in deterministic situation

For starters, let us consider a simplistic and naive deterministic variant of MHE problem statement in order to outline the research landscape. This variant of problem statement ignores possibility of any noise impact and imprecise measurements, empathizing only on insufficiency of data obtained from a single measurement.

Let there be controllable and observable discrete-time linear system, whose evolution can be described as

$$x(k+1) = Ax(k) + Bu(k), \tag{1}$$

$$y(k) = Cx(k), \tag{2}$$

where $x(k)$ is an (unobservable) system's state vector of dimensionalty n for current point of time k; A is an $n \times n$ matrix; $u(k)$ is an input impact (control) of dimensionality r applied at point of time k; B is a $n \times r$ matrix; $y(k)$ is an (observable) measurement vector of dimensionality m at point of time k; C is an $m \times n$ matrix.

For simplicity, A, B and C will be further considered time invariant. Nevertheless, everything discussed here is also applicable when they also change in time; it is only required to know their precise state at each point of time. Corresponding equations can be trivially rewritten for this case.

 ISSN 2786-6491

Formulas (1), (2) describe dependence between values at two consecutive points of time. As only r scalars from $y(k)$ vector can be measured, they do not contain enough data to recover all n scalars of current state $x(k)$ by themselves. Obviously, measurement of at least n scalars is required to recover n scalars. Thus, at least $\lceil \frac{n}{r} \rceil$ measurements of the output vector at different points of time are required, as well as an equation binding states and outputs at different points of time together. The well-known Cauchy formula for linear system (1), (2)

$$x(k+p) = A^p x(k) + \sum_{i=0}^{p-1} A^{p-1-i} Bu(k+i) \tag{3}$$

can be used for this purpose. Thus, having measurements $x(k-s), \ldots, x(k)$, where $s + 1 \geq$ $\geq \lceil \frac{n}{r} \rceil$, and equation (2) applicable for every point of time, we obtain

$$y(k-s+p) = CA^p x(k-s) + \sum_{i=0}^{p-1} CA^{p-1-i} Bu_{k-s+i}, \quad p \in \{0, \ldots, s\}, \tag{4}$$

$$x(k) = A^s x(k-s) + \sum_{i=0}^{s} A^{s-1-i} Bu(k-s+i), \tag{5}$$

where (4) can be conveniently rewritten in trajectorial form as

$$\underbrace{\begin{pmatrix} y(k-s) \\ y(k-s+1) \\ \vdots \\ y(k) \end{pmatrix}}_{y(k,s)} = \underbrace{\begin{pmatrix} C \\ CA \\ \vdots \\ CA^s \end{pmatrix}}_{\Gamma_{s+1}} x(k-s) +$$

$$+ \underbrace{\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ CB & \ddots & & & \vdots \\ CAB & CB & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{s-2}B & CA^{s-3} & \cdots\ CAB & CB & 0 \\ CA^{s-1}B & CA^{s-2} & \cdots\ \cdots & CAB & CB \end{pmatrix}}_{\Phi_{s+1}} \underbrace{\begin{pmatrix} u(k-s) \\ u(k-s+1) \\ \vdots \\ u(k) \end{pmatrix}}_{u(k,s)}. \tag{6}$$

Here $y(k,s)$ is an observation trajectory and $u(k,s)$ is a control trajectory on the historic horizon, Γ_{s+1} is a $m \cdot (s+1) \times n$ observability matrix, Φ_{s+1} is a $m \cdot (s+1) \times r$ historic controllablity matrix. $x(k-s)$ is considered to be unknown variable and everything else in (6) is considered to be known.

Having any n rows of equation (6) with linear-independent corresponding rows of Γ_{s+1} we can trivially recover initial state $x(k-s)$ of this historic horizon in this overly-idealized example. Thus, having (5), we can fast-forward system's state and recover current state.

While this particular approach is useless in practice, more useful models and state recovery approaches discussed further will be formed as its different modifications and complications. Such narration approach was chosen in order to be as explicit as possible about how various forms of nondeterminism were introduced in models.

2. Classic linear least square estimation

The linear least square (LLS) algorithm is the most obvious, well-known and broadly used approach to find estimation of a vector in a overdefined system of linear equations with some kind of noise applied to variables.

This approach implies that an (unknown) noise with some known covariance matrix is additively applied to all output measurements from considered system of linear equations. In order to conform with this presupposition of the method, the model of linear system (1), (2) should be transformed into

$$x(k+1) = Ax(k) + Bu(k), \tag{7}$$

$$y(k) = Cx(k) + \delta(k), \tag{8}$$

where $\delta(k)$ is unknown measurement noise.

The Cauchy formula (3) written for the deterministic linear system model remains unchanged in this case, while relation (4) for historic measurements transforms into

$$\begin{aligned} y(k-s+p) =& CA^p x(k-s) + \\ &+ \sum_{i=0}^{p-1} CA^{p-1-i} B u_{k-s+i} + \delta(k-s+p), \quad p \in \{0,\ldots,s\}. \end{aligned} \tag{9}$$

Equation (6) thus becomes

$$\begin{aligned} \underbrace{\begin{pmatrix} C \\ CA \\ \vdots \\ CA^s \end{pmatrix}}_{\Gamma_{s+1}} x(k-s) = & \underbrace{\underbrace{\begin{pmatrix} y(k-s) \\ y(k-s+1) \\ \vdots \\ y(k) \end{pmatrix}}_{y(k,s)} - \underbrace{\begin{pmatrix} \delta(k-s) \\ \delta(k-s+1) \\ \vdots \\ \delta(k) \end{pmatrix}}_{\delta(k,s)}}_{\tilde{y}(k,s)} - \\ & - \underbrace{\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ CB & \ddots & & & \vdots \\ CAB & CB & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{s-2}B & CA^{s-3} & \cdots\ CAB & CB & 0 \\ CA^{s-1}B & CA^{s-2} & \cdots & CAB & CB \end{pmatrix}}_{\Phi_{s+1}} \underbrace{\begin{pmatrix} u(k-s) \\ u(k-s+1) \\ \vdots \\ u(k) \end{pmatrix}}_{u(k,s)}, \end{aligned} \tag{10}$$

where $y(k,s)$ is a measured observation on historic horizon, $\delta(k,s)$ is and unknown measurement noise and $\tilde{y}(k,s)$ is an (unknown) unnoised observation. Like in previous model, initial state $x(k-s)$ is considered unknown. Here, the well-known linear least squares (LLS) method allows to estimate $x(k-s)$ as following:

$$x_{\text{est.}}(k-s) = \left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Big(y(k,s) - \Phi_{s+1} u(k,s)\Big), \tag{11}$$

where $\Omega(s)$ is a covariance matrix of the errors $\delta(k,s)$, if we know it and want to apply the generalized variant of linear least squares (GLLS). Otherwise, $\Omega(s)$ is often considered to be an identity matrix, which turns (11) into the ordinary LLS method. This method is computationally cheap if we can precompute the generalized pseudoinverse matrix $\left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s)$.

3. Estimatior stability and regularization

Stability of estimators is rarely (if ever) discussed in literature on state estimation and control. Nevertheless, it is important if we want to discuss systematically such things as regularization in state estimators or impact of system's structural properties on estimation precision.

If we discuss regularization, we need a consistent way to measure degree of estimation's disturbance which we are trying to avoid in this way. We also need to understand at cost of forsaking what information we achieve this.

Structural properties of the considered system define limits of estimation precision which is possible to achieve for a particular system. It becomes apparent if we notice that we are forced to use past measurements to reconstruct a present system's state and thus any estimator regardless of its design must somehow (explicitly or implicitly) reconstruct state transformations inside the system.

The most obvious and well-known example of such limits imposed on estimators by system's structure is the notion of linear system's observability. It states that the state $x(\cdot)$ of a linear system (1), (2) can not be estimated completely if $\operatorname{rank}\Gamma_{n-1}$ is less than dimensionality of the state. If we consider a noised linear system like in (7), (8), these limitations go beyond that. In some cases past system's states do not impact its present state significantly enough to derive sufficient information to estimate it with required precision considering noise levels affecting the system. It is important to understand that such cases exist and that this problem inherently can not be solved by improving estimator's design. An example of this kind of limitation will be shown further by example of the generalized linear least squares (GLLS) estimator.

3.1. Estimator stability appraisal by example of GLLS. While contents of this subsection can be considered trivial, empathizing on the notion of problem's stability, its origins, properties and consequences is required for completeness of further discussion.

Here we will follow a traditional approach to measure problem's stability in terms of condition number. As system's state change in time (while statistical properties of noise typically are not), the absolute condition number [13, p. 90]

$$\kappa = \lim_{\varepsilon \to 0} \sup_{\|\Delta x\| \leq \varepsilon} \frac{\|\Delta f\|}{\|\Delta x\|}, \tag{12}$$

seems to be the most appropriate measure (in contrast to the relative condition number). Here Δx is a disturbance of the input data, $\Delta f = f(x+\Delta x) - f(x)$ is a disturbance of solver's result, where f is the problem's solver (in our case — estimator) and x is its undisturbed (i. e. theoretical precise) input data.

If Δf depends linearly on Δx (like in LLS and GLLS), i. e. there is a matrix ∂f such that $\Delta f(\Delta x) = \partial f \cdot \Delta x$, the absolute condition number becomes

$$\kappa = \sup_{\Delta x} \frac{\|\partial f \cdot \Delta x\|}{\|\Delta x\|} = \|\partial f\| \tag{13}$$

by definition of the matrix norm.

In regard to state estimators, the estimator's condition number suggests us how bad will be the impact of noises affecting a system on state estimation. In particular, with the Cauchy–Bunyakovsky–Schwarz inequality we can show that it limits estimation's error norm (don't confuse with variance) as following:

$$\begin{aligned} \mathbb{E}\|\Delta f\| =& \mathbb{E}\left(\frac{\|\Delta f\|}{\|\Delta x\|} \cdot \|\Delta x\|\right) \leq \\ \leq& \sqrt{\mathbb{E}\left(\frac{\|\Delta f\|}{\|\Delta x\|}\right)^2 \cdot \mathbb{E}\|\Delta x\|^2} \leq \\ \leq& \sqrt{\sup_{\Delta x}\left(\frac{\|\Delta f\|}{\|\Delta x\|}\right)^2 \cdot \mathbb{E}\|\Delta x\|^2} = \\ =& \kappa \cdot \sqrt{\mathbb{E}\|\Delta x\|^2}. \end{aligned} \tag{14}$$

Having (13), the absolute condition number of the GLLS estimator for a past state $x(k-s)$ which considers historic horizon of length s (i. e. last $s+1$ measurements $y(k-s), \ldots, y(k)$) can be calculated as following.

From (10), (11) we get

$$x_{\text{est.}}(k-s) = \left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \left(\Gamma_{s+1} x(k-s) + \delta(k,s)\right) =$$
$$= x(k-s) + \left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \delta(k,s), \tag{15}$$

$$\underbrace{x_{\text{est.}}(k-s) - x(k-s)}_{\Delta f_{\text{GLLS}(s,k-s)}} = \underbrace{\left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s)}_{\partial f_{\text{GLLS}(s,k-s)}} \underbrace{\delta(k,s)}_{\Delta x_{\text{GLLS}(s)}}. \tag{16}$$

Thus, absolute condition number for this estimator is

$$\kappa_{\text{GLLS}(s,k-s)} = \left\| \left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \right\|. \tag{17}$$

For LLS (where $\Omega^{-1}(s)$ is the identity matrix) it becomes

$$\kappa_{\text{LLS}(s,k-s)} = \|\Gamma_{s+1}^{+}\|, \tag{18}$$

where $\Gamma_{s+1}^{+} := (\Gamma_{s+1}^{\mathsf{T}} \Gamma_{s+1})^{-1} \Gamma_{s+1}^{\mathsf{T}}$ is the pseudoinverse of Γ_{s+1}. If we consider the Euclidean norm, we get

$$\kappa_{\text{LLS}(s,k-s)} = 1/\sigma_n, \tag{19}$$

where σ_n is the smallest singular value of Γ_{s+1}. The absolute condition number of the LLS estimator is also the smallest possible absolute condition number of GLLS (at least in terms of the Euclidean norm). Let us show this explicitly.

Let there be a singular value decomposition (SVD) of Γ_{s+1} and its compact form:

$$\Gamma_{s+1} = PQR^{\mathsf{T}} = \underbrace{\left(P_1 \,|\, P_2\right)}_{P} \underbrace{\left(\frac{Q_1}{0}\right)}_{Q} R^{\mathsf{T}} = P_1 Q_1 R^{\mathsf{T}}. \tag{20}$$

Here P is an $m \cdot (s+1) \times m \cdot (s+1)$ orthonormal matrix, R is an $n \times n$ orthogonal matrix and Q is $m \cdot (s+1) \times n$ with singular values (in descending order) on its main diagonal. Correspondingly, P_1 consists of the first n columns of P, P_2 contains the rest of its columns and Q_1 is square $n \times n$ part of Q. With this we can transform the generalized pseudoinverse matrix ∂f as following:

$$\partial f_{\text{GLLS}(s,k-s)} = \left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) =$$
$$= \left(RQ_1 P_1^{\mathsf{T}} \Omega^{-1}(s) P_1 Q_1 R^{\mathsf{T}}\right)^{-1} RQ_1 P_1^{\mathsf{T}} \Omega^{-1}(s) =$$
$$= RQ_1^{-1} \left(P_1^{\mathsf{T}} \Omega^{-1}(s) P_1\right)^{-1} \underbrace{Q_1^{-1} R^{\mathsf{T}} R Q_1}_{\text{identity matrix}} P_1^{\mathsf{T}} \Omega^{-1}(s) =$$
$$= RQ_1^{-1} \left(P_1^{\mathsf{T}} \Omega^{-1}(s) P_1\right)^{-1} P_1^{\mathsf{T}} \Omega^{-1}(s). \tag{21}$$

From (16) and (21) we get

$$x_{\text{est.}}(k-s) - x(k-s) = RQ_1^{-1} \left(P_1^{\mathsf{T}} \Omega^{-1}(s) P_1\right)^{-1} P_1^{\mathsf{T}} \Omega^{-1}(s) \delta(k,s) \tag{22}$$

and thus we can transform the formula (13) for the absolute condition number of GLLS

ISSN 2786-6491

like this:

$$\kappa_{\text{GLLS}(s,k-s)} = \sup_{\delta(k,s)} \frac{\|x_{\text{est.}}(k-s) - x(k-s)\|_2}{\|\delta(k,s)\|_2} =$$

$$= \sup_{\delta(k,s)} \frac{\|RQ_1^{-1}\left(P_1^\intercal \Omega^{-1}(s) P_1\right)^{-1} P_1^\intercal \Omega^{-1}(s)\delta(k,s)\|_2}{\|\delta(k,s)\|_2}. \tag{23}$$

With substitution

$$\delta(k,s) = \left(P_1 \,|\, \Omega(s)P_2\right)\begin{pmatrix}\varepsilon_1 \\ \hline \varepsilon_2\end{pmatrix} = P_1\varepsilon_1 + \Omega(s)P_2\varepsilon_2, \tag{24}$$

where $\varepsilon_1 \in \mathbb{R}^n$ and $\varepsilon_2 \in \mathbb{R}^{m\cdot(s+1)-n}$, we can transform it further into this:

$$\kappa_{\text{GLLS}(s,k-s)} = \sup_{\varepsilon_1,\varepsilon_2} \frac{\left\|RQ_1^{-1}\left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} P_1^\intercal \Omega^{-1}(s)\left(P_1 \,|\, \Omega(s)P_2\right)\begin{pmatrix}\varepsilon_1 \\ \hline \varepsilon_2\end{pmatrix}\right\|_2}{\left\|\left(P_1 \,|\, \Omega(s)P_2\right)\begin{pmatrix}\varepsilon_1 \\ \hline \varepsilon_2\end{pmatrix}\right\|_2}. \tag{25}$$

Hence, it can be significantly simplified as following:

$$\left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} P_1^\intercal \Omega^{-1}(s)\left(P_1 \,|\, \Omega(s)P_2\right)\begin{pmatrix}\varepsilon_1 \\ \hline \varepsilon_2\end{pmatrix} =$$

$$= \left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} P_1^\intercal \Omega^{-1}(s)\Big(P_1\varepsilon_1 + \Omega(s)P_2\varepsilon_2\Big) =$$

$$= \underbrace{\left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} P_1^\intercal \Omega^{-1}(s)P_1}_{\text{identity matrix}}\,\varepsilon_1 + \left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} P_1^\intercal \underbrace{\Omega^{-1}(s)\Omega(s)}_{\text{identity matrix}} P_2\varepsilon_2 =$$

$$= \varepsilon_1 + \left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} \underbrace{P_1^\intercal P_2}_{0}\,\varepsilon_2 = \varepsilon_1. \tag{26}$$

And finally, considering that R and P consist of ortonormal columns, we get

$$\kappa_{\text{GLLS}(s,k-s)} = \sup_{\varepsilon_1,\varepsilon_2} \frac{\|RQ_1^{-1}\varepsilon_1\|_2}{\|P_1\varepsilon_1 + \Omega(s)P_2\varepsilon_2\|_2} \geq$$

$$\geq \sup_{\varepsilon_1} \frac{\|RQ_1^{-1}\varepsilon_1\|_2}{\|P_1\varepsilon_1\|_2} = \sup_{\varepsilon_1} \frac{\|Q_1^{-1}\varepsilon_1\|_2}{\|\varepsilon_1\|_2} = \underbrace{1/\sigma_n}_{\kappa_{\text{LLS}(s,k-s)}}. \tag{27}$$

From results obtained above we should notice that condition numbers of LLS and GLLS appeared to be properties of not these algorithms themselves, but of a system they are applied to. Thus, we confirmed that for LLS and GLLS estimators system's properties define by themselves how well its state can be estimated.

3.2. Regularization in GLLS. As we have seen previously, the largest precision loss is related to the smallest singular values of the observation matrix Γ_{s+1}. Thus, it is tempting deal with it by artificially getting rid of the smallest its singular values. Let us show, what happens if we do this and what do we loose with this modification of GLLS.

Having (11), (21) we can express the GLLS estimator as

$$x_{\text{est.}}(k-s) = RQ_1^{-1}\left(P_1^\intercal \Omega^{-1}(s)P_1\right)^{-1} P_1^\intercal \Omega^{-1}(s)\Big(y(k,s) - \Phi_{s+1}u(k,s)\Big). \tag{28}$$

If we zero-out inverses of some smallest observation matrix's singular values in construction of Q_1^{-1} like proposed above, we obtain following regularized estimator:

$$x_{\text{reg. est.}}(k-s) = R\left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array}\right)\Big(P_1^\intercal \Omega^{-1}(s) P_1\Big)^{-1} P_1^\intercal \Omega^{-1}(s) \times$$
$$\times \Big(y(k,s) - \Phi_{s+1} u(k,s)\Big). \tag{29}$$

Here p is the number of singular values we decided to leave as-is and $\sigma_1, \cdots, \sigma_p$ are p largest singular values of Γ_{s+1}.

If we apply (10), like we did in the previous section, we get

$$x_{\text{reg. est.}}(k-s) = R\left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array}\right)\Big(P_1^\intercal \Omega^{-1}(s) P_1\Big)^{-1} P_1^\intercal \Omega^{-1}(s) \times$$
$$\times \Big(\underbrace{P_1 Q_1 R^\intercal}_{\Gamma_{s+1}} x(k-s) + \delta(k,s)\Big) =$$
$$= R\left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array}\right) Q_1 R^\intercal x(k-s) +$$
$$+ R\left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array}\right)\Big(P_1^\intercal \Omega^{-1}(s) P_1\Big)^{-1} P_1^\intercal \Omega^{-1}(s) \delta(k,s) =$$
$$= R\left(\begin{array}{ccc|c} 1 & & & \\ & \ddots^{p \text{ times}} & & 0 \\ & & 1 & \\ \hline & 0 & & 0 \end{array}\right) R^\intercal x(k-s) +$$
$$+ R\left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array}\right)\Big(P_1^\intercal \Omega^{-1}(s) P_1\Big)^{-1} P_1^\intercal \Omega^{-1}(s) \delta(k,s). \tag{30}$$

As we can see, we obtained an estimation $x_{\text{reg. est.}}(k-s)$ of state's projection

$$x_{\text{pr.}}(k-s) = R\left(\begin{array}{ccc|c} 1 & & & \\ & \ddots^{p \text{ times}} & & 0 \\ & & 1 & \\ \hline & 0 & & 0 \end{array}\right) R^\intercal x(k-s), \tag{31}$$

but not of the state $x(k-s)$ itself. This way we forsake recovering a certain state subspace for which signal-noise ratio is too small to consider an estimation usable. In other words, measurements $y(k,s)$, loosely speaking, do not contain enough information about this state subspace, and thus we decided not even to try to estimate it.

Absolute condition number of $x_{\text{reg. est.}}(k-s)$ can be calculated the same way, as of $x_{\text{est.}}(k-s)$ in the previous section. This way we get

 ISSN 2786-6491

$$\kappa_{\text{reg. GLLS}(s,k-s)} = \sup_{\delta(k,s)} \frac{\|x_{\text{reg. est.}}(k-s) - x_{\text{pr.}}(k-s)\|_2}{\|\delta(k,s)\|_2} =$$

$$= \left\| R \left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array} \right) \left(P_1^\intercal \Omega^{-1}(s) P_1 \right)^{-1} P_1^\intercal \Omega^{-1}(s) \right\|_2 =$$

$$= \left\| \left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array} \right) \left(P_1^\intercal \Omega^{-1}(s) P_1 \right)^{-1} P_1^\intercal \Omega^{-1}(s) \right\|_2 \geq$$

$$\geq 1/\sigma_p = \kappa_{\text{reg. LLS}(s,k-s)}. \qquad (32)$$

3.3. Choosing a horizon length and regularization degree. While regularization approach described in the previous subsection can be used as a last resort, we can notice that inability to estimate properly a part of the state-space is due to the lack of information contained in observations on the history horizon we have chosen. Thus, extending considered history horizon may improve estimation by giving additional information.

Absolute conditional numbers of GLLS and its regularized variants for different horizon lengths give us indispensable information regarding what precision we get in which case and thus can be used to make informed judgement about reasonable history horizon length and/or a linear sub-space of the state-space we consider being impossible to estimate. In simplified case of the ordinary LLS (i. e. when $\Omega(s)$ is the identity matrix for all s) the absolute condition numbers are exactly the inverses of singular values of observability matrices for corresponding horizon lengths.

Fig. 1 displays how these singular values change with increasing horizon length for three example systems. Vertical scales on this figure is logarithmic. Typically (i. e. in all cases we have seen), singular values of system's observability matrices are increasing with the historic horizon length. But as we can see on Fig. 1, *a*, *c*, a singular value may not grow indefinitely and asymptotically approach a certain limit value instead.

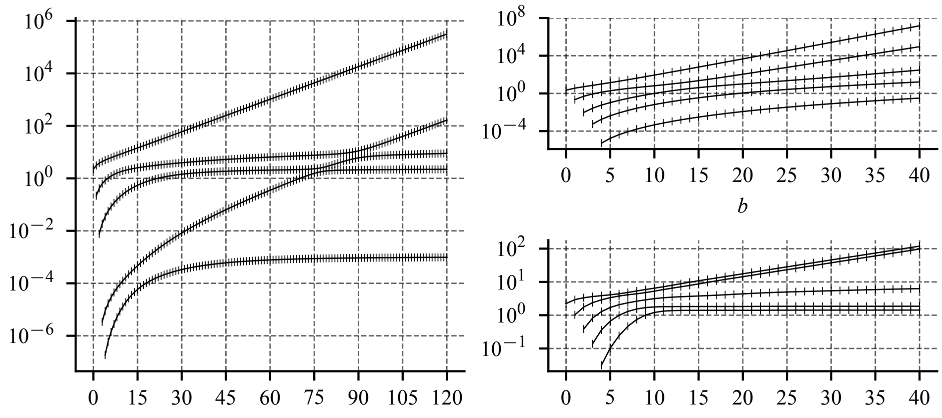

Fig. 1

There are cases when all observability matrice's singular values are limited, when part of them are limited (like on Fig. 1, a, c) and also cases when growth of all of them is unlimited (like on Fig. 1, b). There are also some convoluted cases like in Fig. 1, a.

In order to make a decision, we need to define precisely what does we consider to be the precision of the estimator in general (contrary to a precision of a particular estimation this estimator made). If we use the robust control approach, we suppose that there is a maximum value μ such that in the system (7), (8)

$$\forall k \ \|\delta(k)\|_2 \leq \mu \tag{33}$$

and our objective is to minimize worst possible estimation error under this condition. Then we conclude that

$$\|\delta(k,s)\|_2 \leq \sqrt{s} \cdot \mu \tag{34}$$

and

$$\|x_{\text{est.}}(k-s) - x(k-s)\|_2 \leq \sqrt{s} \cdot \mu \cdot \kappa_{\text{GLLS}(s,k-s)}, \tag{35}$$

$$\|x_{\text{reg. est.}}(k-s) - x(k-s)\|_2 \leq \sqrt{s} \cdot \mu \cdot \kappa_{\text{reg. GLLS}(s,k-s)}. \tag{36}$$

For ordinary LLS it becomes

$$\|x_{\text{est.}}(k-s) - x(k-s)\|_2 \leq \sqrt{s} \cdot \mu \,/\, \sigma_n(\Gamma_{s+1}), \tag{37}$$

$$\|x_{\text{reg. est.}}(k-s) - x(k-s)\|_2 \leq \sqrt{s} \cdot \mu \,/\, \sigma_p(\Gamma_{s+1}), \tag{38}$$

where $\sigma_n(\Gamma_{s+1})$ and $\sigma_p(\Gamma_{s+1})$ are the n-th (the smallest) and the p-th singular values of Γ_{s+1} (which is used to estimate $x(k-s)$).

If our objective is instead to minimize the expected estimation error (i. e. to make best estimations in most cases), we should consider (14). For GLLS it becomes

$$\mathbb{E}\|x_{\text{est.}}(k-s) - x(k-s)\|_2 \leq \sqrt{\mathbb{E}\|\delta(k,s)\|_2^2} \cdot \kappa_{\text{GLLS}(s,k-s)}, \tag{39}$$

$$\mathbb{E}\|x_{\text{reg. est.}}(k-s) - x(k-s)\|_2 \leq \sqrt{\mathbb{E}\|\delta(k,s)\|_2^2} \cdot \kappa_{\text{reg. GLLS}(s,k-s)}. \tag{40}$$

$\sqrt{\mathbb{E}\|\delta(k,s)\|_2^2}$ increases with horizon length growth. In particular, if scalar elements of $\delta(k,s)$ are independent and have the same distribution, then it is proportional to $\sqrt{s \cdot m}$.

As we can see, in both cases our notions of estimator's general precision are proportional to $\sqrt{s} \cdot \kappa_{\text{GLLS}(s,k-s)}$ (or $\sqrt{s} \cdot \kappa_{\text{reg. GLLS}(s,k-s)}$ for regularized estimator) and if we consider only ordinary LLS, it becomes inversely proportional to $\frac{\sigma_n(\Gamma_{s+1})}{\sqrt{s}}$ (or $\frac{\sigma_p(\Gamma_{s+1})}{\sqrt{s}}$ for regularized estimator). Fig. 2 displays this index for the same systems as on Fig. 1. Vertical scales here are also logarithmic.

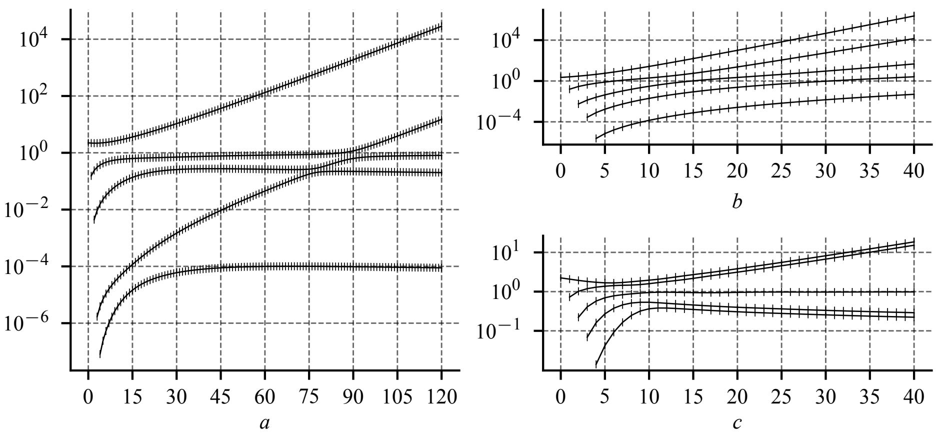

Fig. 2

If index $\frac{\sigma_n(\Gamma_{s+1})}{\sqrt{s}}$ grows indefinitely, like on Fig. 2, *b*, the only things which limit estimation precision are available history of measurements, computational resources and possible numerical problems caused by operating with huge matrices in floating point arithmetic. If there exists $\max_s \frac{\sigma_n(\Gamma_{s+1})}{\sqrt{s}}$, like on Fig. 2, a, c, it is reasonable to consider the corresponding horizon length as the best one. If we, for example, decide for system a, that it is not possible to meaningfully estimate full state, we can discard several smallest singular values. In this case, considering that we discarded $n-p$ singluar values, we should use horizon length corresponding to $\max_s \frac{\sigma_p(\Gamma_{s+1})}{\sqrt{s}}$ instead. Take note that the best horizon length for regularized estimator in most cases is different from one for non-regularized estimator.

We should empathise here that R in the singular value decomposition (20) would be different for different historic horizon lengths s, and thus we can not just simply estimate different components of the state space using different horizon lengths and combine results afterwards.

3.4. Stability and regularization of current state estimation. Careful reader may notice, that while considered estimators are intended to reconstruct system's state in the past, what is actually important to control the system is its current state. Of course, having an estimation of a past state and subsequent controls for system (7), (8), we can fast-forward this estimation like this:

$$x_{\text{est.}}(k) = A^s x_{\text{est.}}(k-s) + \sum_{i=0}^{s-1} A^{s-1-i} B u_{k-s+i}. \tag{41}$$

This solves the problem, but now we need to know an absolute condition number for current state's estimation to make an educated choice of considered historic horizon length s and regarding regularization.

By substituting (3) from (41) we get

$$x_{\text{est.}}(k) - x(k) = A^s \big(x_{\text{est.}}(k-s) - x(k-s)\big) \tag{42}$$

and by substituting (16) into it we obtain

$$\underbrace{x_{\text{est.}}(k) - x(k)}_{\Delta f_{\text{GLLS}(s,k)}} = A^s \left(\Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1}\right)^{-1} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \delta(k,s) =$$

$$= \underbrace{\left((A^{-s})^{\mathsf{T}} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1} A^{-s}\right)^{-1} (A^{-s})^{\mathsf{T}} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s)}_{\partial f_{\text{GLLS}(s,k)}} \underbrace{\delta(k,s)}_{\Delta x_{\text{GLLS}(s)}}. \tag{43}$$

Having this and a singular value decomposition

$$\Gamma_{s+1} A^{-s} = P' Q' R'^{\mathsf{T}} = \underbrace{\left(P'_1 \,|\, P'_2\right)}_{P'} \underbrace{\left(\frac{Q'_1}{0}\right)}_{Q'} R'^{\mathsf{T}} = P'_1 Q'_1 R'^{\mathsf{T}}. \tag{44}$$

we can repeat inferences from two previous subsections and get similar results.

For instance, the absolute condition numbers (implying the Euclidean norm) of the GLLS and ordinary LLS estimators for current state $x(k)$ which consider historic horizon of length s (i. e. last $s+1$ measurements $y(k-s), \ldots, y(k)$) would be

$$\kappa_{\text{GLLS}(s,k)} = \left\| \left((A^{-s})^{\mathsf{T}} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \Gamma_{s+1} A^{-s}\right)^{-1} (A^{-s})^{\mathsf{T}} \Gamma_{s+1}^{\mathsf{T}} \Omega^{-1}(s) \right\|_2 \geq$$

$$\geq 1/\sigma'_n = \kappa_{\text{LLS}(s,k)}, \tag{45}$$

where is the smallest singular value of $\Gamma_{s+1} A^{-s}$.

From (11), (41), (44) we get

$$
\begin{aligned}
x_{\text{est.}}(k) = {} & R' Q_1'^{-1} \left(P_1'^{\mathsf{T}} \Omega^{-1}(s) P_1' \right)^{-1} P_1'^{\mathsf{T}} \Omega^{-1}(s) \Big(y(k,s) - \Phi_{s+1} u(k,s) \Big) + \\
& + \sum_{i=0}^{s-1} A^{s-1-i} B u_{k-s+i},
\end{aligned} \tag{46}
$$

which allows us to do its regularization

$$
\begin{aligned}
x_{\text{reg. est.}}(k) = {} & R' \left(\begin{array}{ccc|c} \sigma_1'^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p'^{-1} & \\ \hline & 0 & & 0 \end{array} \right) \left(P_1'^{\mathsf{T}} \Omega^{-1}(s) P_1' \right)^{-1} P_1'^{\mathsf{T}} \Omega^{-1}(s) \times \\
& \times \Big(y(k,s) - \Phi_{s+1} u(k,s) \Big) + \\
& + R' \left(\begin{array}{ccc|c} 1 & & & \\ & \overset{p \text{ times}}{\ddots} & & 0 \\ & & 1 & \\ \hline & 0 & & 0 \end{array} \right) R'^{\mathsf{T}} \sum_{i=0}^{s-1} A^{s-1-i} B u_{k-s+i},
\end{aligned} \tag{47}
$$

where p is the number of singular values we decided to leave as-is and $\sigma_1', \ldots, \sigma_p'$ are p largest singular values of $\Gamma_{s+1} A^{-s}$.

Like in the previous subsection we can demonstrate, that

$$
\begin{aligned}
x_{\text{reg. est.}}(k) = {} & \underbrace{R' \left(\begin{array}{ccc|c} 1 & & & \\ & \overset{p \text{ times}}{\ddots} & & 0 \\ & & 1 & \\ \hline & 0 & & 0 \end{array} \right) R'^{\mathsf{T}} x(k)}_{x_{\text{pr.}}(k)} + \\
& + R' \left(\begin{array}{ccc|c} \sigma_1'^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p'^{-1} & \\ \hline & 0 & & 0 \end{array} \right) \left(P_1'^{\mathsf{T}} \Omega^{-1}(s) P_1' \right)^{-1} P_1'^{\mathsf{T}} \Omega^{-1}(s) \delta(k,s)
\end{aligned} \tag{48}
$$

and thus

$$
\begin{aligned}
\kappa_{\text{reg. GLLS}(s,k)} = {} & \left\| \left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_p^{-1} & \\ \hline & 0 & & 0 \end{array} \right) \left(P_1'^{\mathsf{T}} \Omega^{-1}(s) P_1' \right)^{-1} P_1'^{\mathsf{T}} \Omega^{-1}(s) \right\|_2 \geq \\
\geq {} & 1/\sigma_p' = \kappa_{\text{reg. LLS}(s,k)}.
\end{aligned} \tag{49}
$$

For some reason, singular values of Γ_{s+1} and $\Gamma_{s+1} A^{-s}$ appeared to be the same (or at least numerically almost the same, considering imprecision of floating point calculations) for all systems tested by authors. This empirical fact should be taken with care and requires further investigation.

4. Quadratic cone state estimators

LLS is a well-known estimator which properties are already deeply researched. As we have demonstrated previously, it can be applied to linear systems like (7), (8) for state

ISSN 2786-6491

estimation. The main presupposition of (7), (8) is that noises are impacting only state measurement, but this is not always the case. For example, a system may be disturbed at the state transition stage instead. This section is dedicated to describing how the GLLS estimator can be modified for such cases.

4.1. Quadratic cone programming form of LLS. The estimation made by GLLS (11) is actually a solution of the following (made-up) problem:

$$\min_{x_{\text{est.}}(k-s)} \sum_{p=0}^{s} \left\| \Omega^{-\frac{1}{2}} \left(y(k-s+p) - y_{\text{est.}}(k-s+p,\, x_{\text{est.}}(k-s)) \right) \right\|_2 \tag{50}$$

considering

$$\begin{aligned} y_{\text{est.}}(k-s+p \,|\, x_{\text{est.}}(k-s)) :=& CA^{p} x_{\text{est.}}(k-s) + \\ &+ \sum_{i=0}^{p-1} CA^{p-1-i} Bu(k-s+i), \quad p \in \{0, \ldots, s\}, \end{aligned} \tag{51}$$

where Ω is a covariance matrix of measurement noise $\delta(k)$ in (8) at each step k, and thus $\Omega^{-\frac{1}{2}}$ is a hermitian square root of its inverse.

Formulating the optimization problem this way is motivated by hypotheses that measurements are where the noise impacts the system. But, obviously, this is not always the case. In order to get some insights on how to modify this problem for other cases, let us reformulate it as a quadratic cone program

$$\min \begin{pmatrix} x_{\text{est.}}(k-s) \\ \delta_{\text{est.}}(k-s) \\ \vdots \\ \delta_{\text{est.}}(k) \end{pmatrix}^{\mathsf{T}} \left(\begin{array}{c|c} \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{n \text{ times}} & \\ \hline & \underbrace{\begin{matrix} \Omega^{-1} & & \\ & \ddots & \\ & & \Omega^{-1} \end{matrix}}_{s+1 \text{ times}} \end{array} \right) \begin{pmatrix} x_{\text{est.}}(k-s) \\ \delta_{\text{est.}}(k-s) \\ \vdots \\ \delta_{\text{est.}}(k) \end{pmatrix} \tag{52}$$

considering

$$\left(\begin{array}{c|cccc} C & \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{r \text{ times}} & & & \\ \hline CA & & \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{r \text{ times}} & & \\ \hline \vdots & & & \ddots & \\ \hline CA^{s} & & & & \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{r \text{ times}} \end{array} \right) \begin{pmatrix} x_{\text{est.}}(k-s) \\ \delta_{\text{est.}}(k-s) \\ \vdots \\ \delta_{\text{est.}}(k) \end{pmatrix} =$$

$$\Gamma_{s+1} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{s} \end{pmatrix}$$

$$= \begin{pmatrix} y(k-s) \\ y(k-s+1) - CBu(k-s) \\ \vdots \\ y(k) - \sum\limits_{i=0}^{s-1} CA^{s-1-i} Bu(k-s+i) \end{pmatrix}, \tag{53}$$

where $\delta_{\text{est.}}(k-s+p) := y(k-s+p) - y_{\text{est.}}(k-s+p,\, x_{\text{est.}}(k-s))$ for $p = \{0, \ldots, s\}$.

After careful consideration, it appears that highlighted parts of (52), (53) are originating not in the system's structure, but in the presupposition about noise properties. Thus, other presupposition would lead us to modification of these parts of the optimization problem producing state estimation. Also take note that the observability matrix Γ_{s+1} is an invariant part of the optimization problem which should proliferate to other modifications.

4.2. Estimation based on noised state model. Having template from previous section, let us design a similar optimization problem for the case when state transition is noised and measurements are precise. A linear system corresponding to this particular presupposition would look like

$$x(k+1) = Ax(k) + Bu(k) + \xi(k), \tag{54}$$

$$y(k) = Cx(k), \tag{55}$$

where $\xi(k)$ is an (unknown) random noise with covariance matrix Ω. The Cauchy formula for system (54), (55) would be

$$x(k+p) = A^p x(k) + \sum_{i=0}^{p-1} A^{p-1-i} Bu(k+i) + \sum_{i=0}^{p-1} A^{p-1-i} \xi(k+i). \tag{56}$$

Like in previous models, we can derive following relations for the historic horizon:

$$\begin{aligned} y(k-s+p) = CA^p x(k-s) + &\\ &+ \sum_{i=0}^{p-1} CA^{p-1-i} Bu(k-s+i) + \\ &+ \sum_{i=0}^{p-1} CA^{p-1-i} \xi(k-s+i), \qquad p \in \{0, \dots, s\}, \end{aligned} \tag{57}$$

$$\text{x(k)} = A^s x(k-s) + \sum_{i=0}^{p-1} A^{p-1-i} Bu(k-s+i) + \sum_{i=0}^{p-1} A^{p-1-i} \xi(k-s+i). \tag{58}$$

These two equations correspond to (4), (5) in the deterministic model and (9), (5) for the LLS model.

The equation (57) in trajectorial form would be

$$\begin{aligned} &\underbrace{\begin{pmatrix} y(k-s) \\ y(k-s+1) \\ \vdots \\ y(k) \end{pmatrix}}_{y(k,s)} = \underbrace{\begin{pmatrix} C \\ CA \\ \vdots \\ CA^s \end{pmatrix}}_{\Gamma_{s+1}} x(k-s) + \\ &+ \underbrace{\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ CB & \ddots & & & \vdots \\ CAB & CB & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{s-2}B & CA^{s-3}B & \cdots CAB & CB & 0 \\ CA^{s-1}B & CA^{s-2}B & \cdots & CAB & CB \end{pmatrix}}_{\Phi_{s+1}} \underbrace{\begin{pmatrix} u(k-s) \\ u(k-s+1) \\ \vdots \\ u(k) \end{pmatrix}}_{u(k,s)} + \\ &+ \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ C & \ddots & & & \vdots \\ CA & C & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{s-2} & CA^{s-3} & \cdots CA & C & 0 \\ CA^{s-1} & CA^{s-2} & \cdots & CA & C \end{pmatrix} \begin{pmatrix} \xi(k-s) \\ \xi(k-s+1) \\ \vdots \\ \xi(k) \end{pmatrix}. \end{aligned} \tag{59}$$

ISSN 2786-6491

Let us apply the approach used in the design of the LLS method in this situation. Like in LLS, we will build an optimization problem whose objective function minimizes estimated values of noises while considering known relations on historic horizon originating from system's dynamics. This way we obtain following optimization problem:

$$\min \sum_{p=0}^{s-1} \|\Omega^{-\frac{1}{2}} \xi_{\text{est.}}(k-s+p)\|_2 \tag{60}$$

considering

$$y_{\text{est.}}(k-s+p\,|\,x_{\text{est.}}(k-s), \xi_{\text{est.}}(k-s), \ldots, \xi_{\text{est.}}(k-s+p-1)) = y(k+s-p), \tag{61}$$

where

$$y_{\text{est.}}(k-s+p\,|\,x_{\text{est.}}(k-s), \xi_{\text{est.}}(k-s), \ldots, \xi_{\text{est.}}(k-s+p-1)) =$$

$$= CA^p x_{\text{est.}}(k-s) + \sum_{i=0}^{p-1} CA^{p-1-i} Bu(k-s+i) + \sum_{i=0}^{p-1} CA^{p-1-i} \xi_{\text{est.}}(k-s+i). \tag{62}$$

If we transform it to the quadratic cone programming form, we obtain

$$\min \begin{pmatrix} x_{\text{est.}}(k-s) \\ \xi_{\text{est.}}(k-s) \\ \vdots \\ \xi_{\text{est.}}(k-1) \end{pmatrix}^{\mathsf{T}} \left(\begin{array}{c|c} \overbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}^{n \text{ times}} & \\ \hline & \overbrace{\begin{matrix} \Omega^{-1} & & \\ & \ddots & \\ & & \Omega^{-1} \end{matrix}}^{s \text{ times}} \end{array}\right) \begin{pmatrix} x_{\text{est.}}(k-s) \\ \xi_{\text{est.}}(k-s) \\ \vdots \\ \xi_{\text{est.}}(k-1) \end{pmatrix} \tag{63}$$

considering

$$\left(\begin{array}{c|ccccc} C & 0 & \cdots & \cdots & \cdots & 0 \\ CA & C & \ddots & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ CA^{s-1} & CA^{s-2} & \cdots & CA & C & 0 \\ CA^s & CA^{s-1} & \cdots & \cdots & CA & C \end{array}\right) \begin{pmatrix} x_{\text{est.}}(k-s) \\ \xi_{\text{est.}}(k-s) \\ \vdots \\ \xi_{\text{est.}}(k-1) \end{pmatrix} =$$

$$\underbrace{\phantom{CA^{s-1}}}_{\Gamma_{s+1}}$$

$$= \begin{pmatrix} y(k-s) & \\ y(k-s+1) - CBu(k-s) & \\ \vdots & \\ y(k) & - \sum_{i=0}^{s-1} CA^{s-1-i} Bu(k-s+i) \end{pmatrix}. \tag{64}$$

Together with (58), this optimization problem allows us to estimate the current state $x(k)$.

As we can see, the problem (63), (64) has the same structure, as (52), (53) for the LLS case, with parts dealing with (highlighted) noise estimations modified. This way we can build a broad class of state estimators corresponding to different presuppositions about system's design and noises affecting it. The core computational primitive of this estimator class would be quadratic cone programming solving, which allows us to use the readily-available CVXOPT solver [14].

5. Computational experiments

5.1. Precision of the LLS estimator compared to its theoretical limit. Fig. 3 shows results of computation experiments made using the same model systems as on previous figures. During each experiment evolution of a system (7), (8) was simulated for 40 steps.

Experiments with each model system started from the same corresponding initial state. Independent noises $\delta(\cdot)$ with independent elements were applied during simulation. Afterwards multiple estimations of the last state using the LLS method were done considering different historic horizons. Then estimations were compared with system's actual state. There were 100 000 such experiments done for each model system.

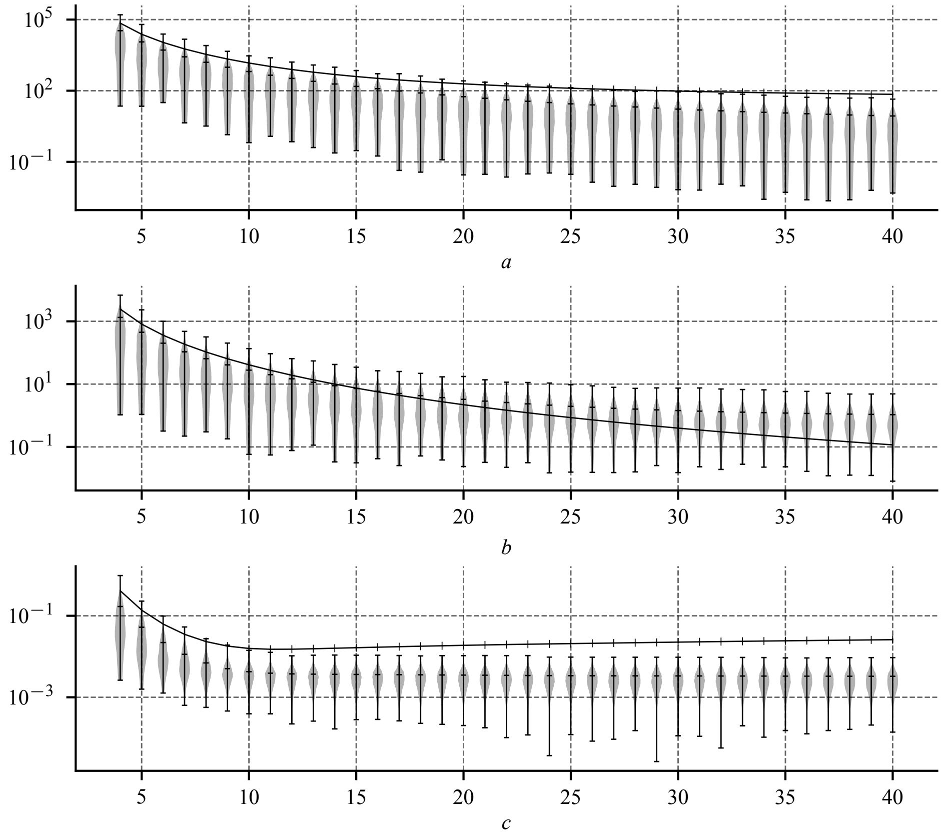

Fig. 3

Violins on the plot show density of estimation precisions obtained during experiments for corresponding horizon lengths on abscissa. In other words, horizon length 4 on abscissa corresponds to estimations made using five measurements $y(k-4), \ldots, y(k)$. Precisions here were calculated as $\|x_{\text{est.}}(k) - x(k)\|_2$. Minimum, maximum and mean of them are marked on each violin.

Values of $\kappa_{\text{LLS}(s,k)} \cdot \sqrt{\mathbb{E}\|\delta(k,s)\|_2^2}$ for corresponding systems and horizon lengths were also plotted for comparison.

As we can see on Fig. 3, *c*, the error limiting inequality (14) can become rather conservative for long horizons. At the same time, an estimator implementation when may fail to make use of theoretically predicted precision improvement on longer horizons because of errors originating from floating point computations, as we can see on Fig. 3, *b*. For comparison, mean norm of system's actual state in experiments on Fig. 3, *b* is approximately $1.47 \cdot 10^7$, which explains the abnormal precision loss.

It is worth to note that while noises $\delta(\cdot)$ used in experiments on Fig. 3 had normal distributions, usage of noises with continuous uniform distribution instead produced similar results.

5.2. Comparison between modified and LLS estimators' performance with system having nosed state. Like in previous subsection, an ensemble of 100 000 experiments was done. Each individual system evolved like in (54), (55) for 16 steps, had the

 ISSN 2786-6491

same matrices and normally distributed transition noises ξ. For instance, eigenvalues of their matrix A was roughly 3.057, $-0.374 \pm 1.130i$, -0.600, 1.095 and singular values are roughly $3.735, 3.166, 1.828, 0.636, 0.206$; i. e. systems were unstable.

Fig. 4 depicts precisions of the final state estimation made with the modified estimator from the Subsection 4.2 (left half-violins) and the LLS estimator (right half-violins). Again, the ordinate has logarithmic scale. Minimum, maximum and mean precisions are marked on violins. Also, mean final state norm of the whole ensemble is depicted as a dash-dotted horizontal line.

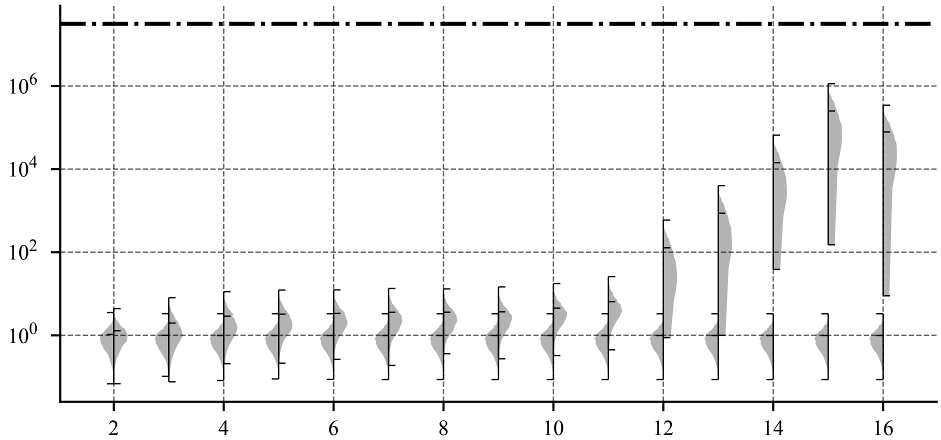

Fig. 4

There are two observations we can make from these results. The first one is that there are indeed such systems with noised state transition for which the modified estimator gives better performance. At the same time we must acknowledge that there are cases when the two estimators have comparable precisions. We associate this with the fact that a system in form (54), (55) can be represented as a system in form (7), (8) and vice versa, but noises transformed this way will obviously have different statistical distribution. In particular, if the system is unstable, like in experiments on Fig. 4, it would amplify past transition noises. In addition, these noises in transformed form will no longer be independent. In some cases it does not affect operation of the LLS estimator, in other ones it does significantly.

Another observation we can make is that the inconsistency between the presupposition about system's design used in the estimator and the system's actual design affects the estimation precision for smaller historical horizon lengths and starts to affect it significantly for longer horizons. It hints us about the source of the frequent rule of thumb to prefer using shorter horizons in time series processing: inconsistencies between presuppositions and reality are (often) less apparent on smaller amounts of data.

Conclusion

The classical GLLS estimator is convenient for estimating state of linear systems with noised indirect measurements. We know how to calculate its absolute condition number for different amounts of considered historic state measurements. Thus we can determine if it is possible to get an estimation with meaningful precision for a particular system and we can choose horizon length which will allow us to get the best results.

Even if it is impossible to get an estimation of the whole system's state with required precision, the regularized variant of the GLLS estimator allows us at least to estimate a state's projection on a subspace, which would have a better precision.

At the same time, it (once again) became apparent that each estimator's design is deeply tied to some hypothesis about target system's construction. In particular, the GLLS estimator's design supposes that state measurement is the only point at which the target linear system is affected by noises. If, for example, the system's state itself is affected at each step, like in (54), (55), the precision of the GLLS estimator may be worse than of one specifically tied to this kind of system. Such precision loss becomes more apparent with horizon length growth. That is why we suggest the following rule of thumb: if estimator's precision worsens too dramatically with increasing amount of source data, it is worth to check other hypotheses about system's design.

Changing considered system design would require development of a new estimator specific to it. Representing the classic GLLS estimator as a quadratic cone programming problem gives us various opportunities to tweak it for this purpose. We developed an estimator tied to system's design represented as (54), (55) as an example of such modification. In practice, a particular considered linear system design should originate from a certain real-world problem, and the purpose of this example is to propose a method of developing specific estimators for non-standard cases. Of course, every new estimator created this way would need a separate research regarding its precision and condition number.

В.Ф. Губарев, М.Д. Міщенко

ІНТЕРВАЛЬНИЙ ОЦІНЮВАЧ СТАНУ ДЛЯ ЛІНІЙНИХ СИСТЕМ З ВІДОМОЮ СТРУКТУРОЮ

Губарев Вячеслав Федорович

Інститут космічних досліджень НАН України та ДКА України, м. Київ,
v.f.gubarev@gmail.com

Міщенко Михайло Дмитрович

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»,
Навчально-науковий інститут прикладного системного аналізу, м. Київ,
mdmisch@firemail.cc, mdmisch@protonmail.com

Керування системою, стан якої не є спостережним безпосередньо, є розповсюдженою задачею. Натомість наявні непрямі, неповні і зашумлені вимірювання стану. У таких випадках фільтр Калмана є загальноприйнятим і класичним підходом до оцінки стану лінійних систем по непрямим вимірюванням. Він рекурсивний і тому опосередковано приймає до уваги усю історію вимірювань. Ми досліджуємо альтернативний підхід: оцінку, виходячи з вимірювань на обмеженому історичному горизонті. У статті спершу обговорюється використання узагальненого методу найменших квадратів (УМНК) щодо цієї задачі, а також умови, при яких доцільно використовувати цей метод. Для випадків, коли він не підходить, пропонуємо спосіб представлення оцінювача за УМНК як задачі квадратичного програмування на конусі, що дає можливість створювати його модифікації, підлаштовані під різноманітні нестандартні конструкції лінійних систем. У статті також досліджено різні властивості і поведінку оцінювача, побудованого за УМНК та модифікаціями цього методу. Зокрема, цілком очікуваним є те, що оцінювачі демонструють різну точність при різній кількості використаних вимірювань. Тому було досліджено застосування абсолютного числа обумовленості оцінювача на базі УМНК до вибору оптимальної довжини горизонту. Було продемонстровано, як абсолютне число обумовленості, будучи жорстким обмеженням точності оцінювання, також обмежує і математичне сподівання норми помилки. Вибір найкращої довжини горизонту було описано з обох цих точок зору. Для ситуацій, коли найкраща можлива точність оцінювання все ще не є достатньою, запропоновано метод регуляризації. Досліджено його переваги та недоліки, а також те, як робити поінформований вибір стосовно ступеня регуляризації. Теоретичні результати перевірено шляхом обчислювальних експериментів.

Ключові слова: лінійна система, оцінка стану, обмежений історичний горизонт вимірювань, метод найменших квадратів, квадратичне програмування на конусі.

REFERENCES

1. Андреев Ю.Н. Управление конечномерными линейными объектами. Москва : Наука, 1976. 424 с.
2. Isermann R., Münchhof M. State and parameter estimation by Kalman filtering. *Identification of dynamic systems: An introduction with applications.* Berlin, Heidelberg : Springer Berlin Heidelberg, 2011. P. 539–551. ISBN: 978-3-540-78879-9. DOI: 10.1007/978-3-540-78879-9_21.
3. Raïssi T., Ramdani N., Candau Y. Set membership state and parameter estimation for systems described by nonlinear differential equations. *Automatica.* 2004. Vol. 40, N 10. P. 1771–1777. DOI: 10.1016/j.automatica.2004.05.006. URL: https://www.sciencedirect.com/science/article/pii/S0005109804001529.
4. Черноусько Ф.Л. Оценивание фазового состояния динамических систем. Метод эллипсоидов. Москва : Наука, 1988. 320 с. ISBN: 5-02-013899-1.
5. Volosov V.V., Tyutyunnik L.I. Development and analysis of robust algorithms for guaranteed ellipsoidal estimation of the state of multidimensional linear discrete dynamic systems. Part 1. *Journal of Automation and Information Sciences.* 2000. Vol. 32, N 3. P. 37–46. DOI: 10.1615/JAutomatInfScien.v32.i3.50.
6. Kuntsevich V.M., Volosov V.V. Ellipsoidal and interval estimation of state vectors for families of linear and nonlinear discrete-time dynamic systems. *Cybernetics and Systems Analysis.* 2015. Jan. Vol. 51, N 1. P. 64–73. DOI: 10.1007/s10559-015-9698-9.
7. Volosov V.V., Shevchenko V.N. Regularized ellipsoidal filter for the state of discrete dynamical systems using averaged current measurements at a sliding interval. *Journal of Automation and Information Sciences.* 2020. Vol. 52, N 9. P. 1–11. DOI: 10.1615/JAutomatInfScien.v52.i9.10.
8. Camacho E.F., Bordons C. Model predictive control. Advanced textbooks in control and signal processing. 2 ed. Springer London, 2007. XXII, 405 p. ISBN: 978-1-85233-694-3. DOI: 10.1007/978-0-85729-398-5. URL: https://link.springer.com/book/10.1007/978-0-85729-398-5.
9. Wu Z., Stoustrup J., Jorgensen J.B. Moving horizon control and estimation of livestock ventilation systems and indoor climate. *IFAC Proceedings Volumes.* 2008. Vol. 41, N 2. P. 6039–6044. 17th IFAC World Congress. DOI: 10.3182/20080706-5-KR-1001.01019. URL: https://www.sciencedirect.com/science/article/pii/S1474667016399116.
10. Muske K.R., Rawlings J.B., Lee J.H. Receding horizon recursive state estimation. *1993 American Control Conference.* San Francisco : IEEE. 1993. P. 900–904. DOI: 10.23919/ACC.1993.4792993.
11. Tenny M., Rawlings J. Efficient moving horizon estimation and nonlinear model predictive control. *Proceedings of the 2002 American Control Conference (IEEE Cat. No.CH37301).* Anchorage : IEEE. 2002. Vol. 6. P. 4475–4480. DOI: 10.1109/ACC.2002.1025355.
12. State estimation for systems subjected to bounded uncertainty using moving horizon approach / Gubarev V.F., Shevchenko V.N., Zhukov A.O., and Gummel A.V. *IFAC Proceedings Volumes.* 2009. Vol. 42, N 10. P. 910–915. 15th IFAC Symposium on System Identification. DOI: 10.3182/20090706-3-FR-2004.00151. URL: https://www.sciencedirect.com/science/article/pii/S1474667016387651.
13. Trefethen L.N., Bau III D. Numerical linear algebra. Philadelphia : Society of Industral and Applied Mathematics, 1997. XII, 361 p. ISBN: 978-0-89871-361-9.
14. Andersen M., Dahl J., Vandenberghe L. CVXOPT: Python software for convex optimization : 2023. URL: https://cvxopt.org/ (online; accessed: 2023-07-25).

Submitted 01.08.2023