Universität Regensburg
Mathematik

State Estimation analysed as Inverse Problem

Luise Blank

Preprint Nr. 17/2005

State Estimation analysed as Inverse Problem

Luise Blank *

Keywords: state estimation, inverse problem, observability measure, optimal control

Abstract: Monitoring dynamical processes requires the estimation of the entire state, which is only partly accessible by measurements. Most quantities must be determined via model based state estimation. Since in general only noisy data are given, state estimation yields an ill-posed inverse problem. Observability guarantees a unique least squares solution. While well-posedness as well as observability is a qualitative behaviour, the quantitative behaviour can be described using the concept of condition numbers. They depend, like stability, crucially on the chosen norms. In this context we shortly review on ill-posed problems, observability and conditioning, and introduce, as a quantification, an observability measure based on condition numbers. For the linear case we show the connection to the well known observability Gramian. For state estimation regularization techniques concerning the initial data are commonly applied in addition to the least squares ansatz. However, we show that the least squares formulation is well-posed and avoids otherwise possibly occuring bias. The introduced observability measure gives a lower bound on the conditioning of this problem formulation. Consequently, inspite of well-posedness a low observability measure causes bad conditioning.

Introducing possible model error functions we leave the finite dimensional setting. To analyse in detail the influence of the regularization parameters and of the coefficients of the model, we study, as a start, linear state equations as constraints. Linear problems appear nearly always as a subproblem of the nonlinear case, and, hence, their features we have to face solving for the nonlinear model. We show that state estimation formulated as optimization problem omitting regularization of the initial data leads to a well-posed problem with respect to L_2- and L_∞- disturbances. However, if the introduced measure of observability is low, the condition numbers of the evolving operators can be arbitrarily large. Small disturbances in the L_2-measurement may propagate to large errors in the states. Nevertheless, for the probably in praxis more relevant L_∞-norm perturbations yield errors in the initial data bounded independently of the system matrix. Finally, we draw conclusions and emphasize the issue of the appropriate norms for state estimation.

1 Introduction

In application the state of a process has to be estimated given noisy data over a past time horizon. These data correspond only to a few state functions, so called output functions. The coupling with all remaining states is given by model equations. This inverse problem is in general ill-posed, since the measurements are noisy and the corresponding continuous signals do not fulfill the model equations. Hence, the existence requirement for well-posedness in the sense of Hadamard is violated. Considering the least squares solution the uniqueness is guaranteed by the observability of the system given by the model equation. The third requirement of well-posedness, namely stability, depends crucially on the norms, which are chosen to measure the disturbances. For state estimation stability may be present in some cases. However, as soon as model error functions are introduced this is not any longer true. Additional regularization is required. Assuming now for state estimation a least squares problem formulation with appropriate regularization to guarantee well-posedness, then there arises the next question: how 'well' does the solution behave in case of disturbances? The corresponding question for the system itself is, how 'well' is the system observable, how high is the observability measurement? Both questions ask for the relation of the output error to the input error. Condition numbers is a general formulated mathematical concept for operators answering this question. Hence, we make use of the concept of condition numbers in the context of state estimation, which leads to a new definition of observability measure. Like for the

*NWF I - Mathematik, University of Regensburg, Germany, luise.blank@mathematik.uni-regensburg.de

question of stability it is essential to choose first in which norms one like to measure the errors to decide whether a problem is well- or ill-conditioned.

Alltogether, this paper discusses for state estimation the well-posedness of the regularized least squares problem formulation, the conditioning, i.e. error propagation, and the influencing observability measure. For all three issues the chosen norms play an important role.

The structure of the paper is the following. In section 2 we first shortly resume the definitions of well-posedness, the possibilities to overcome ill-posedness and the concept of condition numbers. Then we study the model equations. We discuss observability and introduce an observability measure based on condition numbers. For linear model equations and the L_2-norm this observability measure depends on the observability Gramian. We also give an example when we face low observability measures. Moreover we show that the least squares problem formulation for linear state estimation without regularization is sufficient to guarantee well-posedness requiring only observability of the system. Regularization of the initial value would yield unnecessary bias. Nevertheless the error propagation depends on the observability measure, so that for a low measure one could think of applying regularization to obtain better conditioning, changing the problem though.

In the following section we extend the model equations linearly by possible model error functions. In addition we include inequality constraints reflecting e.g. safety constraints. The least squares problem formulation, now necessarily regularized with respect to the error functions, gives an optimization problem, for which we state the first order necessary condition. Then we restrict the analysis to linear state equations omitting inequality constraints. They appear usually as subproblems solving the nonlinear problem and their analysis enhances already some of the main features we face also for the nonlinear case. In particular, we study the influence of the stiffness of the state equation, the measure of observability and of the regularization parameter. As in the first part without model error functions we derive well-posedness for the optimization formulation omitting regularization of the initial data with respect to the L_2-norm and with respect to the L_∞-norm. However, while for one state only we see that the problem is well-conditioned with respect to the L_∞-norm, a low observability measure may result into a ill-conditioned problem with respect to the L_2-norm independent of the regularization parameter. The corresponding operator has the same behaviour.

In the last section we draw conclusions and emphasize the issue of the appropriate choice of norms concerning data errors and state errors.

2 Well-posedness, condition number and observability measure

Typically the noisy measurements $\boldsymbol{z}(t_i) \in \mathbb{R}^{n_y}$ at discrete times are preprocessed. Most algorithms are based on the assumption to have an underlying function corresponding to the discrete data, e.g. for many filtering techniques the Fourier transformation is used at some stage. Hence, it is appropriate to assume a preprocessing of the data on a horizon $[t_0, t_0 + H]$ to a still noisy function $\boldsymbol{z} \in L_2([t_0, t_0 + H], \mathbb{R}^{n_y})$. These data $\boldsymbol{z}$ correspond to a few states, called output functions which we denote with $\boldsymbol{y}$. The output functions $\boldsymbol{y}$ are coupled with all states $\boldsymbol{x}$ and their initial values $\boldsymbol{x}_0 = \boldsymbol{x}(t_0) \in \mathbb{R}^{n_x}$ by model equations. Usually the number of output functions n_y is far less than the number of state functions n_x. Given the model equations one can determine from the initial values the states and therefore the output equations. Hence they define an operator $\boldsymbol{K} : \boldsymbol{x}_0 \to \boldsymbol{y}$. In general we have $\boldsymbol{z} \notin \mathcal{R}(\boldsymbol{K})$ (the range of $\boldsymbol{K}$). This violates the first condition of well-posedness in the sense of Hadamard [8, 12]:

2.1 Ill-posed problems, regularization and condition numbers

Definition 1 *Given an operator $\boldsymbol{K} : \boldsymbol{X} \to \boldsymbol{Y}$ where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are normed spaces, then the equation $\boldsymbol{K}\boldsymbol{x} = \boldsymbol{y}$ is well-posed in the sense of Hadamard iff*

1. *Existence: there exists for all $\boldsymbol{y} \in \boldsymbol{Y}$ a solution $\boldsymbol{x} \in \boldsymbol{X}$; ($\boldsymbol{K}$ surjective).*
2. *Uniqueness: there is at most one solution $\boldsymbol{x} \in \boldsymbol{X}$; ($\boldsymbol{K}$ injective).*
3. *Stability: $\boldsymbol{x}$ depends continuously on $\boldsymbol{y}$, i.e. $\|\boldsymbol{K}\boldsymbol{x}_n - \boldsymbol{K}\boldsymbol{x}\|_{\boldsymbol{Y}} \to 0 \Rightarrow \|\boldsymbol{x}_n - \boldsymbol{x}\|_{\boldsymbol{X}} \to 0$; ($\boldsymbol{K}^{-1}$ continuous).*

The equation is ill-posed if one of these properties does not hold.

It is important to specify the spaces as well as the topologies of the spaces, i.e. the norms $\|\cdot\|_{\boldsymbol{X}}$ and $\|\cdot\|_{\boldsymbol{Y}}$. The problem can be well-posed using one set of norms and ill-posed in another set of norms. If the problem is ill-posed there are several remedies, of which we recall only some relevant in our context.

Assume that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are Hilbert-spaces (i.e. there exists a scalar product; e.g. the space L_2) and $\boldsymbol{K} : \boldsymbol{X} \to \boldsymbol{Y}$ is linear and compact then $\boldsymbol{x}$ is called *least-squares solution* (best fit) if $\boldsymbol{x}$ is the solution of $\min_{\boldsymbol{x}\in\boldsymbol{X}} \|\boldsymbol{K}\boldsymbol{x}-\boldsymbol{z}\|_{\boldsymbol{Y}}$. Moreover, it holds: $\boldsymbol{x}$ is the least squares solution if and only if the *normal equation* $\boldsymbol{K}^*\boldsymbol{K}\boldsymbol{x} = \boldsymbol{K}^*\boldsymbol{z}$ holds, where $\boldsymbol{K}^*$ denotes the adjoint operator. In case of a finite-dimensional space $\boldsymbol{X}$ this ansatz overcomes the failure of existence.

Uniqueness is not necessarily an issue in our context since we require observability of the system given by the model equations (see later). However, otherwise one can use the Moore-Penrose inverse, also called *generalized inverse*, which is given as the unique best-approximate solution, i.e. the least squares solution of minimal norm [8], if there exists a least squares solution. For finite-dimensional $\boldsymbol{X}$ the generalized inverse is given by $\boldsymbol{K}^\dagger := (\boldsymbol{K}^*\boldsymbol{K})^{-1}\boldsymbol{K}^*$.

However, the generalized inverse does not overcome the lack of continuity in general. Regularization techniques have to be applied. Here, we can distinguish roughly speaking three kinds of approaches. Tikhonov regularization shifts the spectrum of $\boldsymbol{K}^*\boldsymbol{K}$ and leads to the *regularized generalized inverse* $\boldsymbol{R}_d := (d\boldsymbol{I} + \boldsymbol{K}^*\boldsymbol{K})^{-1}\boldsymbol{K}^*$, which is bounded, with a regularization parameter $d > 0$. Solving $\boldsymbol{R}_d\boldsymbol{x} = \boldsymbol{z}$ is equivalent to the minimization problem

$$\min_{\boldsymbol{x}\in\boldsymbol{X}} \|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{z}\|^2_{\boldsymbol{Y}} + d\|\boldsymbol{x}\|^2_{\boldsymbol{X}}. \tag{1}$$

Regularization by discretization (projection) discretizes $(\boldsymbol{K}^*\boldsymbol{K})^{-1}\boldsymbol{K}^*$ or respectively the corresponding equation, which leads to an operator $\boldsymbol{R}_d$ where d depends on discretization level. Roughly speaking the high frequencies of $(\boldsymbol{K}^*\boldsymbol{K})^{-1}\boldsymbol{K}^*$ are cut off. An application in state estimation can be found i.e. in [3]. Iterative regularization methods solve iteratively $\min \|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{z}\|$ e.g. with a steepest descent method, or respectively solve the normal equation with an iterative fix point scheme, but stop after n iterations. It yields a regularized generalized inverse $\boldsymbol{R}_d$ with d depending on the number of iterations n. The operator $\boldsymbol{R}_d\boldsymbol{K}$ should converge pointwise to the identity for $d \to 0$. Moreover, the choice of the regularization parameter d should give the best compromise between data and regularization error, i.e. let $\|\boldsymbol{z} - \boldsymbol{z}^\delta\| \le \delta$ and $\boldsymbol{x}^{d,\delta} = \boldsymbol{R}_d\boldsymbol{y}^\delta$ then

$$\|\boldsymbol{x}^{d,\delta} - \boldsymbol{x}\| \le \|\boldsymbol{R}_d(\boldsymbol{z}^\delta - \boldsymbol{z})\| + \|\boldsymbol{R}_d\boldsymbol{z} - \boldsymbol{x}\| \le \|\boldsymbol{R}_d\|\delta + \|\boldsymbol{R}_d\boldsymbol{z} - \boldsymbol{x}\|, \tag{2}$$

should be minimal. The first term is called data error and the second regularization error. This is a non trivial task but will not be discussed in this paper, instead we refer to the literature [8, 12, 14].

Now let us assume that $\boldsymbol{K}\boldsymbol{x} = \boldsymbol{y}$ is a well-posed problem. Then $\boldsymbol{K}^{-1}$ exists and is bounded with respect to the chosen norms. That means, the equation is stable, which is a qualitative statement. The mathematical concept of condition number is quantitative. It measures the possible error propagation with respect to the absolute or relative error [7].

Definition 2 *Considering the problem given $\boldsymbol{y}$ determining the solution $\boldsymbol{x}$ of $\boldsymbol{K}\boldsymbol{x} = \boldsymbol{y}$*
1. the absolute normwise condition number is the smallest number $\kappa_{abs}(\boldsymbol{y}) > 0$ such that for $\|\tilde{\boldsymbol{y}}-\boldsymbol{y}\|_{\boldsymbol{Y}} \to 0$

$$\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_{\boldsymbol{X}} = \|\boldsymbol{K}^{-1}\tilde{\boldsymbol{y}} - \boldsymbol{K}^{-1}\boldsymbol{y}\|_{\boldsymbol{X}} \le \kappa_{abs}(\boldsymbol{y})\|\tilde{\boldsymbol{y}} - \boldsymbol{y}\|_{\boldsymbol{Y}} + o\left(\|\tilde{\boldsymbol{y}} - \boldsymbol{y}\|_{\boldsymbol{Y}}\right); \tag{3}$$

2. the relative normwise condition number is the smallest number $\kappa_{rel}(\boldsymbol{y}) > 0$ such that for $\|\tilde{\boldsymbol{y}}-\boldsymbol{y}\|_{\boldsymbol{Y}} \to 0$

$$\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_{\boldsymbol{X}}/\|\boldsymbol{x}\|_{\boldsymbol{X}} \le \kappa_{rel}(\boldsymbol{y})\|\tilde{\boldsymbol{y}} - \boldsymbol{y}\|_{\boldsymbol{Y}}/\|\boldsymbol{y}\|_{\boldsymbol{Y}} + o\left(\|\tilde{\boldsymbol{y}} - \boldsymbol{y}\|_{\boldsymbol{Y}}/\|\boldsymbol{y}\|_{\boldsymbol{Y}}\right). \tag{4}$$

The problem is called well-conditioned if κ is small and ill-conditioned for large κ.
For linear $\boldsymbol{K}$ we have

$$\kappa_{abs}(\boldsymbol{y}) \le \|\boldsymbol{K}^{-1}\|_{\boldsymbol{Y}\to\boldsymbol{X}} \text{ and } \kappa_{rel}(\boldsymbol{y}) \le \|\boldsymbol{K}\|_{\boldsymbol{X}\to\boldsymbol{Y}}\|\boldsymbol{K}^{-1}\|_{\boldsymbol{Y}\to\boldsymbol{X}}. \tag{5}$$

If $\boldsymbol{K}$ is a matrix, the condition number is defined as the latter $cond(\boldsymbol{K}) := \|\boldsymbol{K}\|\|\boldsymbol{K}^{-1}\|$, where commonly the l_2-norms are used.

2.2 Observability measure

For state estimation on the horizon $[t_0, t_0+H]$ the operator $\boldsymbol{K} : \boldsymbol{x}_0 \longmapsto \boldsymbol{y}$ is given by the *model equations*:

State equations:
$$\boldsymbol{G}\dot{\boldsymbol{x}} - \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}) = \boldsymbol{0} \tag{6}$$
$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{7}$$
Output equations:
$$\boldsymbol{y} - \boldsymbol{C}\boldsymbol{x} = \boldsymbol{0} \tag{8}$$

The system (6)-(8) is called *observable*, if for any given $\boldsymbol{u}$ and $\boldsymbol{p}$ the initial state $\boldsymbol{x}_0$ can be uniquely determined from the output $\boldsymbol{y}$ [6, 11]. Hence, $\boldsymbol{K} : \boldsymbol{x}_0 \longmapsto \boldsymbol{y}$ is injective for fixed $\boldsymbol{u}$, $\boldsymbol{p}$ and $\boldsymbol{K}^{-1}$ exists on $\mathcal{R}(\boldsymbol{K})$. The space $\boldsymbol{X}$ is the finite-dimensional space $\mathbb{R}^{n_x}$. Observability is the qualitative behaviour that a difference in the states shall be seen in the outputs. The observability measure shall quantify this statement, hence we consider

$$\|\boldsymbol{y} - \tilde{\boldsymbol{y}}\| \geq c\|\boldsymbol{x}_0 - \tilde{\boldsymbol{x}}_0\| \tag{9}$$

or a relative measurement independent of the scaling

$$\|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|/\|\boldsymbol{y}\| \geq c\|\boldsymbol{x}_0 - \tilde{\boldsymbol{x}}_0\|/\|\boldsymbol{x}_0\|. \tag{10}$$

As larger c as better the observability measure. This suggest the use of the condition number $\kappa = 1/c$ of the problem given $\boldsymbol{y}$ determining the solution of $\boldsymbol{K}\boldsymbol{x}_0 = \boldsymbol{y}$. The evaluation of the conditioning is mentioned also in [1] in preference to the yes/no answer of observability.

Definition 3 *The absolute and the relative measure of observability of $\boldsymbol{x}_0$ are defined as $1/\kappa_{abs}$ and $1/\kappa_{rel}$. The system is called well observable for $\boldsymbol{x}_0$, if $\kappa = 1/c$ is small, and has a low observability measure for large κ.*

For linear model equations the corresponding operator $\boldsymbol{K}$ is affine. Let us first consider linear $\boldsymbol{K}$, i.e. the model equations are linear and $\boldsymbol{p}$ and $\boldsymbol{u} = \boldsymbol{0}$. Without loss of generality we consider in the rest of the paper only the case $t_0 = 0$. Thus we have with

$$\begin{aligned} \dot{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{0}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad \boldsymbol{y} - \boldsymbol{C}\boldsymbol{x} = \boldsymbol{0} & (11)\\ \Rightarrow \boldsymbol{K}\boldsymbol{x}_0 &= \boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{x}_0 & (12)\end{aligned}$$

Choosing now as norm on $\boldsymbol{Y}$ the $L_2([0,H])^{n_y}$-norm we obtain

$$\|\boldsymbol{K}\boldsymbol{x}_0\|_{L_2}^2 = \boldsymbol{x}_0^T \int_0^H (e^{\boldsymbol{A}t})^T \boldsymbol{C}^T \boldsymbol{C} e^{\boldsymbol{A}t}\, dt\, \boldsymbol{x}_0 = \boldsymbol{x}_0^T \mathcal{G}(H)\boldsymbol{x}_0 \tag{13}$$

where the matrix $\mathcal{G}(H) \in \mathbb{R}^{n_x \times n_x}$ is the known finite time observability Gramian (e.g. [5, 11, 13]).

Lemma 1 *Let the system be observable, then:*
a.) The observability Gramian $\mathcal{G}(H) = \int_0^H (e^{\boldsymbol{A}t})^T \boldsymbol{C}^T \boldsymbol{C} e^{\boldsymbol{A}t}\, dt \in \mathbb{R}^{n_x \times n_x}$ is symmetric positive definite, and therefore invertible.
b.) Let $\boldsymbol{v}$ be a real eigenvalue of $\boldsymbol{A}$ to an eigenvalue $\alpha \in \mathbb{R}$. Then $\|\mathcal{G}(H)\|_2$ is large for large α and for a long horizon $[0,H]$, while $\|\mathcal{G}(H)^{-1}\|_2$ is large if $-\alpha$ is large or $\|\boldsymbol{C}\boldsymbol{v}\|_{l_2}$ is small or if the horizon is short.

Proof: we omit the depedency of $\mathcal{G}$ on H in the following.
a.) Symmetry is obvious. Given $\boldsymbol{v} \neq 0$ then $\boldsymbol{y}(t) = \boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{v} \not\equiv \boldsymbol{0}$ since the system is observable. Hence, $\boldsymbol{v}^T\mathcal{G}\boldsymbol{v} = \int_0^H \boldsymbol{y}^T(t)\boldsymbol{y}(t)\, dt = \|y\|_{L_2}^2 > 0$, and $\mathcal{G}$ is positive definite.
b.) Let $\boldsymbol{v}$ and $\boldsymbol{\alpha}$ fulfill the assumption, then

$$\boldsymbol{v}^T\mathcal{G}\boldsymbol{v} = \|e^{\alpha t}\|_{L_2(0,H)}^2 \|\boldsymbol{C}\boldsymbol{v}\|_{l_2}^2 = \frac{e^{2\alpha H} - 1}{2\alpha}\|\boldsymbol{C}\boldsymbol{v}\|_{l_2}^2. \tag{14}$$

With $||\mathcal{G}||_2 = \max_{\boldsymbol{v}\in\mathbb{R}^{n_x}}(\boldsymbol{v}^T\mathcal{G}\boldsymbol{v})/(\boldsymbol{v}^T\boldsymbol{v})$, $\|\mathcal{G}^{-1}\|_2 = \max_{\boldsymbol{v}\in\mathbb{R}^{n_x}}(\boldsymbol{v}^T\boldsymbol{v})/(\boldsymbol{v}^T\mathcal{G}\boldsymbol{v})$ follows the assertion.

q.e.d.

Using the l_2-norm for $\boldsymbol{X} = \mathbb{R}^{n_x}$ it follows for $\boldsymbol{K}$:

$$\begin{aligned} \|\boldsymbol{K}\|_{l_2\to L_2}^2 &= \sup_{\boldsymbol{x}_0\in\mathbb{R}^{n_x}} \frac{\|\boldsymbol{K}\boldsymbol{x}_0\|_{L_2}^2}{\|\boldsymbol{x}_0\|_{l_2}^2} = \sup_{\boldsymbol{x}_0\in\mathbb{R}^{n_x}} \frac{\boldsymbol{x}_0^T\mathcal{G}\boldsymbol{x}_0}{\boldsymbol{x}_0^T\boldsymbol{x}_0} = ||\mathcal{G}||_2 & (15)\\ \|(\boldsymbol{K}\,|_{\mathcal{R}(\boldsymbol{K})})^{-1}\|_{L_2\to l_2}^2 &= \sup_{\boldsymbol{x}_0\in\mathbb{R}^{n_x}} \frac{\|\boldsymbol{x}_0\|_{l_2}^2}{\|\boldsymbol{K}\boldsymbol{x}_0\|_{L_2}^2} = ||\mathcal{G}^{-1}||_2 & (16)\end{aligned}$$

For linear systems with not necessarily $\boldsymbol{p}$ and $\boldsymbol{u} = \boldsymbol{0}$ we need to consider for the condition numbers $\|\boldsymbol{K}\tilde{\boldsymbol{x}}_0 - \boldsymbol{K}\boldsymbol{x}_0\|_{L_2}^2 = (\tilde{\boldsymbol{x}}_0 - \boldsymbol{x}_0)^T\mathcal{G}(\tilde{\boldsymbol{x}}_0 - \boldsymbol{x}_0)$. Hence, having $\sqrt{\|\mathcal{G}^{-1}\|_2}\|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|_{L_2} \geq \|\boldsymbol{x}_0 - \tilde{\boldsymbol{x}}_0\|_{l_2}$ for all $\boldsymbol{x}_0$, we can define for linear systems -like condition numbers for matrices- a observability measure independent of the state $\boldsymbol{x}_0$:

Definition 4 *The absolute, respectively the relative observability measure with respect to the l_2 and L_2-norms for linear systems is given by*

$$1/\sqrt{||\mathcal{G}^{-1}||_2} \quad \textit{respectively} \quad 1/\sqrt{cond(\mathcal{G})} \tag{17}$$

This definition is in agreement with the Gramian based measure in [15], where one considers a infinite horizon. There, several measures are proposed and compared, which are all rather based on the various tests for observability than motivated by error propagation.

Lemma 1 immediately shows that long horizons are better for observability reasons. It also reassures that the eigenvectors of system matrix $\boldsymbol{A}$ should not be close to the null-space of the output matrix $\boldsymbol{C}$. For rapidly decaying systems it is confirmed that it is difficult to determine the initial value exactly, while for the forward problem $\boldsymbol{K}$ the value at the end point is sensitive to $\alpha \gg 0$.

2.3 State estimation as a least squares problem

Going back to the inverse problem of state estimation we obtain

Theorem 1 *For a observable system the problem formulation*

$$\min \|\boldsymbol{y}-\boldsymbol{z}\|_{L_2}^2 + d\|\boldsymbol{x}_0 - \boldsymbol{x}_0^{ref}\|_{l_2}^2 \quad s.t. \quad \dot{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{u},\ \boldsymbol{x}(0) = \boldsymbol{x}_0,\ \boldsymbol{y} - \boldsymbol{C}\boldsymbol{x} = \boldsymbol{0} \quad on\ [0,H] \tag{18}$$

is well-posed for all $d \geq 0$, and the solution is given by

$$\boldsymbol{x}_0 = (\mathcal{G}(H) + d\boldsymbol{I})^{-1}\left[\int_0^H e^{\boldsymbol{A}^T t}\boldsymbol{C}^T\left\{\boldsymbol{z}(t) - \boldsymbol{C}e^{\boldsymbol{A}t}\int_0^t e^{-\boldsymbol{A}s}\boldsymbol{u}(s)\,ds\right\}\,dt + d\boldsymbol{x}_0^{ref}\right]. \tag{19}$$

Regularization of the initial data ($d \neq 0$) is not necessary, and leads to bias if inexact reference values are used.

Proof: Setting $\hat{\boldsymbol{z}} = \boldsymbol{z} - \boldsymbol{C}\int_0^t e^{\boldsymbol{A}(t-s)}\boldsymbol{u}(s)\,ds$ we have

$$\min \|\boldsymbol{y}-\boldsymbol{z}\|_{L_2} + d\|\boldsymbol{x}_0 - \boldsymbol{x}_0^{ref}\|_{l_2}^2 = \min \|\hat{\boldsymbol{y}}-\hat{\boldsymbol{z}}\|_{L_2} + d\|\boldsymbol{x}_0 - \boldsymbol{x}_0^{ref}\|_{l_2}^2 \tag{20}$$

where $\hat{\boldsymbol{y}}$ fulfills the model equations with $\hat{\boldsymbol{u}} = \boldsymbol{0}$. Then, dropping for convenience the $\hat{}$ we have the equivalence of (18) to the normal equation:

$$\Leftrightarrow \min \|\boldsymbol{K}\boldsymbol{x}_0 - \boldsymbol{z}\|_{L_2} + d\|\boldsymbol{x}_0 - \boldsymbol{x}_0^{ref}\|_{l_2}^2 \Leftrightarrow (\boldsymbol{K}^*\boldsymbol{K} + d\boldsymbol{I})\boldsymbol{x}_0 = \boldsymbol{K}^*\boldsymbol{z} + d\boldsymbol{x}_0^{ref} \tag{21}$$

$$\text{with} \quad \boldsymbol{K}^*\boldsymbol{z} = \int_0^H e^{\boldsymbol{A}^T t}\boldsymbol{C}^T\boldsymbol{z}(t)dt, \tag{22}$$

$$\text{since} \quad (\boldsymbol{\xi}_0, \boldsymbol{K}^*\boldsymbol{y})_{l_2} = (\boldsymbol{K}\boldsymbol{\xi}_0, \boldsymbol{y})_{L_2} = \boldsymbol{\xi}_0^T\int_0^H (e^{\boldsymbol{A}t})^T\boldsymbol{C}^T\boldsymbol{y}(t)\,dt. \tag{23}$$

Hence we know $\mathcal{G} = \boldsymbol{K}^*\boldsymbol{K}$ which is invertible for observable systems, and we obtain (19). Moreover, since $\mathcal{G}$ is finite-dimensional it has a bounded inverse. Therefore, (18) is a well-posed problem even for $d = 0$, i.e. without regularizing the initial value. Given a noise free signal $\boldsymbol{z}$, there exists a unique $\boldsymbol{x}_0^{exact}$ s.t. $\boldsymbol{K}\boldsymbol{x}_0^{exact} = \boldsymbol{z}$. We obtain

$$\boldsymbol{x}_0 = \boldsymbol{x}_0^{exact} + (\mathcal{G} + d\boldsymbol{I})^{-1}d(\boldsymbol{x}_0^{ref} - \boldsymbol{x}_0^{exact}) \tag{24}$$

which answers the question of bias. q.e.d.

(Remark: the result concerning the bias is not in contrast to the probabilistic ansatz leading to the Kalman filter, since there one would choose also $d = 0$ for noise free signals [10].)

For $d = 0$, i.e. for the least squares formulation the solution (19) and its extension to time varying systems can be found also in [5], where in addition the differential equation determining the Gramian is derived:

$$\frac{d}{dH}\mathcal{G} = \boldsymbol{A}^T\mathcal{G} + \mathcal{G}\boldsymbol{A} + \boldsymbol{C}^T\boldsymbol{C}, \quad \mathcal{G}(0) = \boldsymbol{0}. \tag{25}$$

There the view point of least squares optimization is taken and the connection to the Kalman filter is made.

While the least squares formulation is well-posed, it is not necessarily well-conditioned.

Corollary 1 *The condition number of (18) with $d = 0$, i.e. of the least squares problem $\boldsymbol{K}^\dagger = \mathcal{G}^{-1}\boldsymbol{K}^*$ with respect to the l_2 and L_2-norms obeys*

$$\kappa_{abs} \geq \|(\boldsymbol{K}\,|_{\mathcal{R}(\boldsymbol{K})})^{-1}\|^2_{L_2\to l_2} = \sqrt{\|\mathcal{G}^{-1}\|_2} \quad \textit{and} \quad \kappa_{rel} \geq \sqrt{cond(\mathcal{G})}. \tag{26}$$

Proof: Let $\boldsymbol{v}$ be the normed eigenvector to the smallest eigenvalue of $\mathcal{G}$, then $\|\mathcal{G}^{-1}\|_2 = \|\mathcal{G}^{-1}\boldsymbol{v}\|_2$. For $\boldsymbol{z}(t) = k\boldsymbol{C}e^{\boldsymbol{A}t}\mathcal{G}^{-1}\boldsymbol{v}$ and $\boldsymbol{u} = \boldsymbol{0}$ we have $\boldsymbol{x}_0 = k\mathcal{G}^{-1}\boldsymbol{v}$. Furthermore, $\|\boldsymbol{z}\|^2_{L_2} = k^2\|\mathcal{G}^{-1}\|_2$ and therefore $\|\boldsymbol{x}_0\| = \|\boldsymbol{z}\|_{L_2}\sqrt{\|\mathcal{G}^{-1}\|_2}$. With (15) and (16) the assertion holds. q.e.d.

Consequently, a low observability measure leads to an ill-conditioned least squares formulation, even though observability provides well-posedness. For example, we may face large error propagation if the assumptions of Lemma 1b.) hold.

3 Inclusion of model error functions

3.1 Optimality conditions

In the following we include linearly possible model error functions $\boldsymbol{w}$ in the model equations. With this step we leave the finite dimensional setting. Considering only the least squares solution does not guarantee stability any longer, as the example of $z = \delta \sin \frac{n}{\delta} t$ for $\dot{x} = ax + w$ shows. Regularization with respect to $\boldsymbol{w}$ is necessary. As regularization parameters we employ now matrices instead of scalars. For the mathematical consideration it is at this point no issue to distinguish the given parameters $\boldsymbol{p}$ and the controls $\boldsymbol{u}$. We summarize them to $\boldsymbol{u}$. Equation (7) for the initial condition can be omitted, since it does not contain any information. Additional inequality constraints reflect safety constraints as well as model verification. Summarized we consider in the following the Tikhonov-type regularized least squares solution of:

$$\min \frac{1}{2}\int_0^H (\boldsymbol{y}-\boldsymbol{z})^T\boldsymbol{Q}(\boldsymbol{y}-\boldsymbol{z}) + \boldsymbol{w}^T\boldsymbol{R_w}\boldsymbol{w}\; dt + \frac{1}{2}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref})^T\boldsymbol{D}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref}) \tag{27}$$

s.t.

$$\boldsymbol{G}\dot{\boldsymbol{x}} - \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{W}\boldsymbol{w} = \boldsymbol{0} \tag{28}$$

$$\boldsymbol{y} - \boldsymbol{C}\boldsymbol{x} = \boldsymbol{0} \tag{29}$$

$$\boldsymbol{c}(\boldsymbol{x},\boldsymbol{u}) \geq \boldsymbol{0}. \tag{30}$$

Obviously, we can substitute $\boldsymbol{y}$ by $\boldsymbol{C}\boldsymbol{x}$ and reduce the system by $\boldsymbol{y}$, the output equations and avoid Lagrange multipliers for these. In addition, setting up the necessary first order equations we see $\boldsymbol{w}$ can be eliminated by the Lagrange multiplier with resp. to the DAE's, namely $\boldsymbol{w} = \boldsymbol{R_w}^{-1}\boldsymbol{W}^T\boldsymbol{\lambda}$. This is a major reduction in size since $\boldsymbol{w}(t)$ may be in $\mathbb{R}^{n_x}$. Defining $\boldsymbol{R} := \boldsymbol{W}\boldsymbol{R_w}^{-1}\boldsymbol{W}^T$ we obtain the following *necessary conditions* (for details see [4], for the linear case without inequality constraints see [10], where also the connection to the Kalman filter is given):

$$\boldsymbol{G}\dot{\boldsymbol{x}} - \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{R}\boldsymbol{\lambda} = \boldsymbol{0} \tag{31}$$

$$-\frac{d}{dt}(\boldsymbol{G}^T\boldsymbol{\lambda}) - \left(\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u})\right)^T\boldsymbol{\lambda} + \boldsymbol{C}^T\boldsymbol{Q}\boldsymbol{C}\boldsymbol{x} + \left(\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{c}(\boldsymbol{x},\boldsymbol{u})\right)^T\boldsymbol{\nu} = \boldsymbol{C}^T\boldsymbol{Q}\boldsymbol{z} \tag{32}$$

$$(\boldsymbol{G}^T\boldsymbol{\lambda})(0) = \boldsymbol{D}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref}) \quad \text{and} \quad (\boldsymbol{G}^T\boldsymbol{\lambda})(H) = \boldsymbol{0} \tag{33}$$

$$\boldsymbol{\nu}^T\boldsymbol{c}(\boldsymbol{x},\boldsymbol{u}) = 0 \quad \boldsymbol{c}(\boldsymbol{x},\boldsymbol{u}) \leq \boldsymbol{0} \quad \boldsymbol{\nu} \geq \boldsymbol{0}. \tag{34}$$

As a result we obtain in case of ODE's as state constraints with no regularization of the initial state, that the model error functions fulfill $\boldsymbol{w} \in H^1$ and $\boldsymbol{w} = 0$ at the end points. If model error functions are present in all state equations, the Lagrange parameter $\boldsymbol{\lambda}$ can be eliminated too with $\boldsymbol{\lambda} = \boldsymbol{R}^{-1}(\boldsymbol{G}\dot{\boldsymbol{x}} - \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}))$. Then, the necessary conditions reduce to a second order DAE system with mixed boundary constraints for the state $\boldsymbol{x}$ only and the equations (34) for the inequality constraints.

3.2 Analysis for linear ODE's

In the solution process for a nonlinear problem with inequality constraints nearly always optimization problems with linear model equations and without inequality constraints (or an equivalent linear equation system) appear as subproblems. Although for the linear case, in particular with regularization of the initial data, efficient methods as the Kalman filter are well established [10, 11], these methods and its extensions cannot be applied or do not perform sufficiently in the presence of nonlinearity and inequality constraints [9]. However, the study of linear model equations without inequality constraints enhances already some of the main features we face for the treatment of the optimization formulation of the nonlinear state estimation. The goal of the study here is to iluminate the influence of the eigenvalues of the system matrix, the influence of the observability measure and the influence of the regularization parameters. As a start we assume possible model error functions in all state equations. We consider the following problem:

$$\min \frac{1}{2}\int_0^H (\boldsymbol{C}\boldsymbol{x}-\boldsymbol{z})^T\boldsymbol{Q}(\boldsymbol{C}\boldsymbol{x}-\boldsymbol{z})+\boldsymbol{w}^T\boldsymbol{R_w}\boldsymbol{w}\,dt+\frac{1}{2}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref})^T\boldsymbol{D}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref}) \quad (35)$$

$$\text{s.t.} \qquad \dot{\boldsymbol{x}}-\boldsymbol{A}\boldsymbol{x}-\boldsymbol{w}=\boldsymbol{u}. \quad (36)$$

A detailed analysis of this problem and its results can be found in [4], where they are presented for $H=1$ and $\boldsymbol{D}=0$. With a few modifications they can be extended to any $H>0$. In this paper we summarize some of the main results. To study the properties we choose one of the following three equivalent formulation, which derive from elimination of the error function $\boldsymbol{w}$ using the state equation or from the necessary conditions eliminating the Lagrange parameter $\boldsymbol{\lambda}$. The boundary value problem and its weak formulation are not only necessary but also sufficient condition, which is shown later (Theorem 4):

Optimization problem:

$$\min_{\boldsymbol{x}\in H^1}\left\{\|\boldsymbol{Q}^{1/2}(\boldsymbol{C}\boldsymbol{x}-\boldsymbol{z})\|_{L_2}^2+\|\boldsymbol{R}^{-1/2}(\dot{\boldsymbol{x}}-\boldsymbol{A}\boldsymbol{x}-\boldsymbol{u})\|_{L_2}^2+\|\boldsymbol{D}^{1/2}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref})\|_{l_2}^2\right\}. \quad (37)$$

Second order BVP:

$$-\boldsymbol{R}^{-1}\ddot{\boldsymbol{x}}+\left(\boldsymbol{R}^{-1}\boldsymbol{A}-\boldsymbol{A}^T\boldsymbol{R}^{-1}\right)\dot{\boldsymbol{x}}+\left(\boldsymbol{A}^T\boldsymbol{R}^{-1}\boldsymbol{A}+\boldsymbol{R}^{-1}\dot{\boldsymbol{A}}+\boldsymbol{C}^T\boldsymbol{Q}\boldsymbol{C}\right)\boldsymbol{x}=\boldsymbol{C}^T\boldsymbol{Q}\boldsymbol{z}-\boldsymbol{R}^{-1}\dot{\boldsymbol{u}}-\boldsymbol{A}^T\boldsymbol{R}^{-1}\boldsymbol{u} \quad (38)$$

with boundary conditions $\dot{\boldsymbol{x}}(0)-(\boldsymbol{A}+\boldsymbol{R}\boldsymbol{D})\boldsymbol{x}(0)=\boldsymbol{u}(0)-\boldsymbol{R}\boldsymbol{D}\boldsymbol{x}_0^{ref}$ and $\dot{\boldsymbol{x}}(H)-\boldsymbol{A}\boldsymbol{x}(H)=\boldsymbol{u}(H)$.

Weak Formulation:

$$\langle\dot{\boldsymbol{\zeta}}-\boldsymbol{A}\boldsymbol{\zeta},\boldsymbol{R}^{-1}(\dot{\boldsymbol{x}}-\boldsymbol{A}\boldsymbol{x})\rangle+\langle\boldsymbol{C}\boldsymbol{\zeta},\boldsymbol{Q}\boldsymbol{C}\boldsymbol{x}\rangle+\boldsymbol{\zeta}^T(0)\boldsymbol{D}(\boldsymbol{x}(0)-\boldsymbol{x}_0^{ref}) \quad (39)$$

$$=\langle\boldsymbol{C}\boldsymbol{\zeta},\boldsymbol{Q}\boldsymbol{z}\rangle+\langle\dot{\boldsymbol{\zeta}}-\boldsymbol{A}\boldsymbol{\zeta},\boldsymbol{R}^{-1}\boldsymbol{u}\rangle \quad \text{for all } \boldsymbol{\zeta}\in H^1.$$

In the following we concentrate only on the case $\boldsymbol{D}=\boldsymbol{0}$, i.e. on the least squares formulation without regularization of the initial data. Among other things we show its well-posedness. With regularization of the initial data this can be studied in the framework of Tikhonov regularization.

For one state function only ($\boldsymbol{A}=\alpha,\boldsymbol{C}=\delta,\boldsymbol{Q}=q,\boldsymbol{R}^{-1}=r$) one can use the BVP to derive an explicit formula for the solution. Analysing this solution we obtain the following theorem.

Theorem 2

1. *The regularized problem formulation (35) (with $\boldsymbol{D}=\boldsymbol{0}$), i.e. given z determining x_0 and w, is well-posed.*

2. *Small perturbation of z in the L_2-norm may lead to large error propagation in the initial data x_0 independently of r and q if $-\alpha$ is large.*

3. *For perturbations of z in the L_∞-norm we have bounds for the errors in x_0 and w independently of α.*

For several state functions we use the weak formulation to show well-posedness. As a first step we again study first the case of one state function where observability is not an issue and extend then the result to several state functions. Let us define the symmetric bilinear form $a: H^1(0,H)\times H^1(0,H)\to\mathbb{R}$

$$a(\boldsymbol{\zeta},\boldsymbol{x})=\langle\dot{\boldsymbol{\zeta}}-\boldsymbol{A}\boldsymbol{\zeta},\boldsymbol{R}^{-1}(\dot{\boldsymbol{x}}-\boldsymbol{A}\boldsymbol{x})\rangle+\langle\boldsymbol{C}\boldsymbol{\zeta},\boldsymbol{Q}\boldsymbol{C}\boldsymbol{x}\rangle \quad (40)$$

If $\boldsymbol{A},\boldsymbol{C}\in\mathbb{R}$ then a is positive definite if $\boldsymbol{R}^{-1},\boldsymbol{Q}>0$. Hence it defines an operator $\boldsymbol{\mathcal{S}}: H^1(0,H)\longrightarrow(H^1(0,H))'$ with

$$(\boldsymbol{\xi},\boldsymbol{\mathcal{S}}\boldsymbol{x}):=a(\boldsymbol{\zeta},\boldsymbol{x})$$

and the weak formulation (39) is equivalent to

$$(\boldsymbol{\xi}, \boldsymbol{\mathcal{S}x}) = \langle \boldsymbol{C\zeta}, \boldsymbol{Qz}\rangle + \langle \dot{\boldsymbol{\zeta}} - \boldsymbol{A\zeta}, \boldsymbol{R}^{-1}\boldsymbol{u}\rangle \quad \text{for all } \boldsymbol{\zeta} \in H^1. \tag{41}$$

We would like to remark that, if we use a Galerkin discretization of the optimization problem (35) then the properties of $\mathcal{S}$ govern to a large extend also the numerical method. For example the condition number of the discretization matrix is influenced by the condition number of $\boldsymbol{\mathcal{S}}$. Hence not only for the well-posedness it is of interest to study the properties of $\boldsymbol{\mathcal{S}}$ but also for the numerical solution approaches.

Using methods of functional analysis one can show the following result in the case of one state function:

Theorem 3 *$\boldsymbol{\mathcal{S}}$ is a linear isomorphism and with the real values $\boldsymbol{A} = \alpha, \boldsymbol{C} = \delta, \boldsymbol{Q} = q, \boldsymbol{R}^{-1} = r$ we have*

$$\min(4r, q\delta^2) \leq \quad \|\boldsymbol{\mathcal{S}}\|_{H^1\to(H^1)'} \quad \leq 2r\max(1,\alpha^2) + q\delta^2, \tag{42}$$

$$\frac{c(\alpha)}{q\delta^2} \leq \quad \|\boldsymbol{\mathcal{S}}^{-1}\|_{(H^1)'\to H^1} \quad \leq \max\left\{\frac{2}{r}, \frac{2\alpha^2+1}{q\delta^2}\right\}, \tag{43}$$

$$\textit{with} \qquad c(\alpha) \approx \quad |\alpha|^{3/2}\sqrt{\frac{e^{2H}-1}{2(e^{2H}+1)}} \qquad \textit{for large } |\alpha|. \tag{44}$$

Hence cond($\boldsymbol{\mathcal{S}}$) $= \|\boldsymbol{\mathcal{S}}\|_{H^1\to(H^1)'}\|\boldsymbol{\mathcal{S}}^{-1}\|_{(H^1)'\to H^1}$ is bounded but tends to infinity with $|\alpha| \to \infty$ for fixed regularization parameters r and q.

The exact value of $c(\alpha) := \|\exp(\alpha\cdot)\|_{H^1}/\|\exp(\alpha\cdot)\|_{(H^1)'}$ is

$$c^2(\alpha) = \frac{(1+\alpha^2)(\alpha+1)^2(\alpha-1)^2(e^H - e^{-H})(e^{2\alpha H}-1)}{(2\alpha+1)(\alpha-1)^2(e^H-e^{-H})(e^{2\alpha H}-1) + 4\alpha^3 e^{-H}(e^{\alpha H}-e^H)^2}. \tag{45}$$

Considering several state functions one has to take into account observability to show positive definiteness and herewith continuity and coercivity of a. We shortly sketch this step while refering for the other arguments to [4].

Given $a(\boldsymbol{x},\boldsymbol{x}) = 0$ then $\boldsymbol{x}$ is the solution of the system $\dot{\boldsymbol{x}} - \boldsymbol{Ax} \equiv \boldsymbol{0}$, $\boldsymbol{x}(0) = \boldsymbol{x}_0$ and $(\boldsymbol{y} \equiv)\boldsymbol{Cx} \equiv \boldsymbol{0}$. Since the system is observable $\boldsymbol{y} \equiv \boldsymbol{0}$ yields $\boldsymbol{x}_0 = \boldsymbol{0}$ and consequently $\boldsymbol{x} \equiv \boldsymbol{0}$. Hence, $a(\boldsymbol{x},\boldsymbol{x}) > 0$ for $\boldsymbol{x} \not\equiv 0$. Continuity and coercivity of a yield

Theorem 4 *For any $\boldsymbol{z}, \boldsymbol{u} \in L_2$ the solution $\boldsymbol{x}$ of the weak formulation (39) determines the unique solution of the minimization problem (37).*

Hence for well-posedness only the question of stability has still to be answered. Using the Riesz representation theorem [2] and considering like for one state only and Lemma 1b.) the exponential function we obtain

Theorem 5 *$\boldsymbol{\mathcal{S}} : H^1 \to \left(H^1\right)'$ is bounded and has a bounded inverse. Furthermore:*

$$0 < 2/\|\boldsymbol{R}\| \quad \leq \quad \|\boldsymbol{\mathcal{S}}\|_{H^1\to(H^1)'} \leq 2\|\boldsymbol{R}\|^{-1}\max(1,\|\boldsymbol{A}\|^2) + \|\boldsymbol{C}^T\boldsymbol{QC}\|, \tag{46}$$

$$\max\{c(\alpha)/\|\boldsymbol{C}^T\boldsymbol{QCv}\|_{l_2}\} \quad \leq \quad \|\boldsymbol{\mathcal{S}}^{-1}\|_{(H^1)'\to H^1}. \tag{47}$$

for all $\boldsymbol{v} \in \mathbb{R}^{n_x}$, $\|\boldsymbol{v}\|_{l_2} = 1$ eigenvector of $\boldsymbol{A}$ with real eigenvalue α. The lower bound in (46) is valid if there exists an $\alpha^2 > 1$. Observability guarantees $\boldsymbol{Cv} \neq \boldsymbol{0}$.

As a consequence of Theorem 5 and Lemma 1b.) we have:

Corollary 2 *For any fixed regularization cond($\boldsymbol{\mathcal{S}}$) is large, if there exists an in modulo large real eigenvalue of $\boldsymbol{A}$ or if there exists a real eigenvector $\boldsymbol{v}$ of $\boldsymbol{A}$ which is close to the null space of $\boldsymbol{C}$. If this is the case then the observability measure is low too.*

Nevertheless, the boundedness of the inverse and the compact inbedding $H^1 \hookrightarrow C^0$ yields

Corollary 3 *Linear state estimation formulated as least squares problem*

$$\min \quad \frac{1}{2}\int_0^H (\boldsymbol{y}-\boldsymbol{z})^T\boldsymbol{Q}(\boldsymbol{y}-\boldsymbol{z}) + \boldsymbol{w}^T\boldsymbol{R_w}\boldsymbol{w}\; dt \tag{48}$$

$$\textit{s.t.} \qquad \dot{\boldsymbol{x}} - \boldsymbol{Ax} - \boldsymbol{w} = \boldsymbol{u} \qquad \boldsymbol{y} - \boldsymbol{Cx} = \boldsymbol{0} \tag{49}$$

is well-posed, i.e. $\|\boldsymbol{x} - \boldsymbol{x}^\delta\|_{H^1} \leq c\|\boldsymbol{z} - \boldsymbol{z}^\delta\|_{(H^1)'}$ and, consequently, with a generic constant c

$$\max\{|x_0 - x_0^\delta|, \|\boldsymbol{w} - \boldsymbol{w}^\delta\|_{L_2}\} \leq c\|\boldsymbol{z} - \boldsymbol{z}^\delta\|_{L_2} \quad \textit{and} \quad \|\boldsymbol{x} - \boldsymbol{x}^\delta\|_{C^0} \leq c\|\boldsymbol{z} - \boldsymbol{z}^\delta\|_{L_2} \leq c\|\boldsymbol{z} - \boldsymbol{z}^\delta\|_{L_\infty}. \tag{50}$$

4 Conclusions and questions concerning the appropriate norms

Observability is like well-posedness a qualitative property. Condition numbers quantify the error propagation. We used this concept to define an observability measure. For linear systems we derived the use of the inverse of observability Gramian $\mathcal{G}$. With $\mathcal{G}^{-1}$ one can estimate the minimal difference in the outputs which can be seen for given different initial data. Also we showed that regularization of initial data is not necessary for stability, hence the least squares formulation for state estimation is well-posed. Regularization of initial data would lead to bias. However, a low observability measure leads to an ill-conditioned least squares problem.

Introducing linearly model error functions we leave the finite dimensional setting. For this case we stated the first order necessary condition and reduced them by several variables and equations. Analysing them for linear systems we showed that the least squares problem formulation without regularizing the initial data is well-posed not only with respect to the L_2-norm but also for the L_∞-norm. The operator corresponding to the reduced system of necessary and sufficient conditions is also bounded and has a bounded inverse though may be ill-conditioned depending on the model equations. Moreover, the error propagation with respect to L_2-errors may be large for low observability measures independent of the regularization parameter for the model errors. Considering L_∞-errors the behaviour is different. Then, in case of one state only, we have bounds independent of the stiffness of state equations.

As seen it is fundamental to discuss in which norms the data errors are bounded and what output is of interest. In my opinion one has not only the L_2-norm of the data error bounded, i.e. $\|\boldsymbol{z}^\delta\|_{L_2(t_0,t_0+H)} \leq \delta$, but one can assume $\|\boldsymbol{z}^\delta\|_{L_2(t_0,t_0+H)} \leq \delta\sqrt{H}$ and $\|\boldsymbol{z}^\delta\|_\infty \leq c$. Hence, one has additional information about the error which should be taken into account, and the error would depend on the length of the horizon. The question concerning the outputs depends on the application of state estimation. Which output is of interest should be stated together with the problem formulation. For example, employing state estimation to obtain the current state required for the main issue of controling a process should have the focus on $\boldsymbol{x}(t_0+H)$, the filtered state. Then the L_2-norm of the state on $[t_0, t_0+H]$ is less adequate than measuring the error of $\boldsymbol{x}(t_0+H)$. If one is interested on the state over the whole horizon, also called smoothed state, one may consider a weighted L_2-norm putting more weight on the current state than on the past state. Or, is the L_∞- norm over the whole horizon more adequate than the L_2-norm? The answers of these question do not only influence the theoretical analysis but should also affect the numerical studies. In particular, as soon as adaptivity concerning the underlying discretization grid is introduced it is of greatest importance to know which error shall be finally small to obtain greatest efficency.

References

[1] Allgöwer, F. and Badgwell, T.A. and Qin, J.S. and Rawlings, J.B. and Wright, S.J., “Nonlinear Predictive Control and Moving Horizon Estimation - An Introductory Overview”, *Advances in Control: Highlights of ECC 1999*, ed. Frank, P.M., Springer-Verlag, 391-449, (1999).

[2] Alt, H.W., “Lineare Funktionalanalysis”, *Springer*, Berlin, Heidelberg, (1999).

[3] Binder, T. and Blank, L. and Dahmen, W. and Marquardt, W., “On the Regularization of Dynamic Data Reconciliation Problems”, *Journal of Process Control*, **12**, 557-567, (2002).

[4] Blank, L., “State Estimation without Regularizing the Initial Data”, *Inverse Problems*, **20**,5, 1357-1370, (2004).

[5] Corless, M.J. and Frazho, A.E., “Linear Systems and Control, An Operator Perspective”, *Marcel Dekker, Inc.*, New York, Basel, (2003).

[6] Curtain, R.F. and Zwart, H., “An introduction to infinite-dimensional linear systems theory”, *Springer*, New York, (1995).

[7] Deuflhard, P. and Hohmann, A., “Numerische Mathematik I. Eine algorithmisch orientierte Einführung”, *Walter de Gruyter*, Berlin, NewYork, (2002).

[8] Engl, H.W. and Hanke, M. and Neubauer, A., “Regularization of Inverse Problems”, *Kluwer*, Dordrecht, The Netherlands, (1996).

[9] Haseltine, E.L. and Rawlings, J.B., "Critical evaluation of extended Kalman filtering and moving horizon estimation", *Ind. Eng. Chem. Res.*,**44**,8, 2451-2460, (2005).

[10] Jazwinski, A.H., "Stochastic Processes and Filtering Theory", *Academic Press*, New York, (1970).

[11] Kailath, T., "Linear Systems", *Prentice Hall*, Englewood Cliffs, New Jersey, (1980).

[12] Kirsch, A., "An Introduction to the Mathematical Theory of Inverse Problems", *Springer-Verlag*, (1996).

[13] Muske, K.R. and Edgar, T.F., "Nonlinear state estimation", *Nonlinear Process Control*, eds. Henson, M.A. and Seborg, D.E., 311-711, (1997).

[14] Tikhonov, A.N. and Goncharsky, A.V. and Stepanov, V.V. and Yagola, A.G., "Numerical Methods for the Solution of Ill-posed Problems", *Kluwer Academic Publishers*, New York, (1996).

[15] Waldraff, W. and Dochain, D. and Bourrel, S. and Magnus, A., "On the Use of Observability Measures for Sensor Location in Tubular Reactors", *Journal of Process Control*, **8**, 497-505, (1998).