Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Econometrics 118 (2004) 341–373
JOURNAL OF Econometrics
www.elsevier.com/locate/econbase

Deterministic least squares filtering

J.C. Willems*

Department of Electrical Engineering, University of Leuven, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium

Dedicated to Manfred Deistler on occasion of his sixtieth birthday

Abstract

A deterministic interpretation of the Kalman filtering formulas is given, using the principle of least squares estimation. The observed signal and the to-be-estimated signal are modeled as being generated as outputs of a finite-dimensional linear system driven by an input disturbance. Postulating that the observed signal is generated by the input disturbance that has minimal least squares norm leads to a method of computing an estimate of the to-be-estimated output. The derivation of the resulting filter is carried out in a completely self-contained way. The analogous approach to least squares control is also discussed.

Keywords: Filtering; Least squares estimation; Kalman filtering; Misfit; Latency; Least squares control; Linear systems; Riccati equation

1. Introduction

It is a pleasure and an honor to have been invited to contribute an article to this special issue dedicated to Manfred Deistler on the occasion of his sixtieth birthday. I will address an issue that has been the topic of numerous discussions and exchanges of ideas which I had with Manfred over the years. Namely, the rationale of using a stochastic setting for obtaining algorithms in dynamical system problems and time-series analysis.

The effectiveness in applications of stochastic models for accommodating uncertainty in dynamical systems is beyond question, it appears. Stochastics is being applied,

* Corresponding author. Tel.: +32-1632-1805; fax: +32-1632-1970.
E-mail addresses: jan.willems@esat.kuleuven.ac.be, j.c.willems@math.rug.nl (J.C. Willems).
URLs: http://www.esat.kuleuven.ac.be/∼willems, http://www.math.rug.nl/∼willems

0304-4076/03/$ - see front matter
doi:10.1016/S0304-4076(03)00146-5

routinely and often successfully, in diverse areas ranging from signal processing to communication networks, from financial mathematics to statistical and quantum physics, from medical diagnosis to epidemiology, etc. Unfortunately, the logical basis and the interpretation of such models is not always very convincing. Particularly in time-series analysis, it is often not clear what a stochastic model implies physically, what it claims about reality. Usually, in fact, it is not stated, even remotely explicitly, if probability is to be interpreted as relative frequency, or as degree of belief, or as plausibility. Moreover, while for finite outcome spaces the probabilistic assumptions are often reasonable and acceptable, it is much more difficult to fathom how one could justify the numerical values of the complex quantitative statements that are implicitly implied by a stochastic time-series as a model of reality.

The purpose of this paper is to show that the Kalman filter admits a perfectly satisfactory deterministic least squares formulation. I will argue that this approach is eminently reasonable, and circumvents the sticky modeling assumptions that are unavoidable in the stochastic approach. The aim of the paper is to give a tutorial and self-contained derivation of these formulas for continuous-time systems. In order to achieve this, I have included at the end brief sections which review standard material from linear system theory and about the Riccati equation.

The Kalman filtering algorithm originally appeared in Kalman (1960a) for discrete-time systems and in Kalman and Bucy (1961) for continuous-time systems. The continuous-time filter is sometimes referred to as the *Kalman-Bucy filter*. Since then, these very important algorithms have been covered in numerous textbooks (for example, Kwakernaak and Sivan, 1972), special issues (for example, Athans, 1971), and reprint volumes (for example, Kailath, 1977). The filtering formulas, and many of their spin-offs, are prominently present in Başar (2001). The Kalman filter actually deals with the same problem as Wiener-Kolmogorov filtering (Wiener, 1949; Kolmogoroff, 1939), a connection that was often alluded to in the original work of Kalman. However, the Kalman filter offers a solution that is far superior to Wiener-Kolmogorov filtering, especially in view of the recursive nature of the algorithm and the effective use that is made of the Riccati equation as a method for computing the filter coefficients. For these and other reasons, the Kalman filter is an algorithm that is of historical significance indeed.

Originally, before this research area became dominated by the intricacies of stochastic calculus, the least squares aspects of the Kalman filter were emphasized. It is very well-known that the Kalman filter is the least squares estimator in a *stochastic* sense. Indeed, the Kalman filter gives a recursive formula for the linear estimator that minimizes the expected value of the square of the estimation error. If the processes involved are jointly gaussian, then the optimal linear estimator coincides with the optimal non-linear estimator, the conditional mean estimator, and the maximum likelihood estimator. These connections between the Kalman filtering algorithm and least squares, in a stochastic context, are classical.

In this paper, I will give a purely deterministic interpretation of the Kalman filter. The basic idea is to explain the observations as being generated by the input disturbances of least squares norm. By substituting these least squares disturbances in the system dynamics, an estimate of any related system variable can be obtained. That

this leads to the same formulas as maximum likelihood estimation when these disturbances are assumed to be stochastic and normally distributed is easy to see for static estimators (see Section 3), and for discrete-time Kalman filtering over a finite time interval. However, mathematical difficulties prevent such a straightforward interpretation for infinite-time filtering and for continuous-time systems. This has to do with the properties of white noise and Brownian motion. In particular, while it is intuitively reasonable to claim that realizations of white noise that have small $\mathscr{L}_2$-norm are more likely than those that have large $\mathscr{L}_2$-norm, mathematically, realizations of white noise have with probability one *infinite* $\mathscr{L}_2$-norm on any non-zero length interval, and so this likelihood interpretation is, at best, an informal one.

My claims what regards originality are very modest. As already mentioned, in static estimation problems (see Section 3), it is well-known and easy to see that there is an equivalence between deterministic least squares on the one hand, and conditional expectation for random vectors with gaussian distributions on the other hand. That a similar least squares interpretation is possible for the Kalman filter is certainly part of the *system theory folklore*, and has been so ever since the Kalman filter appeared on the horizon. The paper by Swerling (1971), in fact, deals exactly with this aspect (see also the references in Swerling's paper). There are a number of earlier sources that give a deterministic interpretation of the Kalman filtering formulas, in particular (Mortenson, 1968; Sontag, 1990; Hijab, 1980; Fleming, 1997; McEneaney, 1998). Of course, the derivation of the analogue of the static case promises to be much more involved for the Kalman filter, in view of the dynamical aspects of the problem setting. The purpose of this paper is to present this derivation.

2. Notation

A few words about the mathematical notation that will be used. As usual, $\mathbb{R}$ denotes the real line, $\mathbb{R}_+ = [0, \infty)$, $\mathbb{R}^n =$ the n-dimensional vectors with real components, $\mathbb{R}^{n\times m} =$ the $n \times m$ matrices with real coefficients. $^\top$ denotes transposition. A symmetric matrix $M = M^\top \in \mathbb{R}^{n\times n}$ is said to be *non-negative definite* (*non-positive definite*), denoted $M \geqslant 0$ ($M \leqslant 0$), if $a^\top M a \geqslant 0$ ($a^\top M a \leqslant 0$) for all $a \in \mathbb{R}^n$, and *positive definite* (*negative definite*), denoted $M \succ 0$ ($M \prec 0$), if $a^\top M a > 0$ ($a^\top M a < 0$) for all $0 \neq a \in \mathbb{R}^n$. For $M = M^\top \in \mathbb{R}^{n\times n}$ and $a \in \mathbb{R}^n$, $a^\top M a$ is often denoted as $\|a\|_M^2$. The Euclidean norm (corresponding to $M = I$) of $a \in \mathbb{R}^n$ is denoted by $\|a\|$.

The map f from the set A to the set B is denoted by $f : A \to B$. If f takes the element $a \in A$ to $b \in B$, we write $f : a \mapsto b$, or $f : a \in A \mapsto b \in B$.

Let A be a (finite or infinite) interval in $\mathbb{R}$. $\mathscr{L}_2(A, \mathbb{R}^n)$ denotes the set of all square integrable maps from A to $\mathbb{R}^n$. The $\mathscr{L}_2$-norm of $f \in \mathscr{L}_2(A, \mathbb{R}^n)$, is defined by $\|f\|^2_{\mathscr{L}_2(A,\mathbb{R}^n)} = \int_A \|f(t)\|^2 \, dt$. When A is an infinite interval in $\mathbb{R}$, we denote by $\mathscr{L}_2^{\text{loc}}(A, \mathbb{R}^n)$ ('$\mathscr{L}_2$-local') all $f : A \to \mathbb{R}^n$ such that $\int_{t_0}^{t_1} \|f(t)\|^2 \, dt < \infty$ for all finite intervals $[t_0, t_1] \subset A$. We denote by $\mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$ ('half-line' $\mathscr{L}_2$) the maps $f : \mathbb{R} \to \mathbb{R}^n$ such that $\int_{-\infty}^{T} \|f(t)\|^2 \, dt < \infty$ for all $T \in \mathbb{R}$.

3. Static estimation

Assume that d is a d-dimensional vector of real random variables, whose joint distribution is normal with (for simplicity) zero mean and covariance normalized to the identity. Consider the problem of estimating $z = Hd$ from observing $y = Cd$, with $H \in \mathbb{R}^{z\times d}$ and $C \in \mathbb{R}^{y\times d}$ (fixed, known) matrices. Assume (again for simplicity) that C has rank y. Then it is well-known and elementary to prove that

$$\hat{z} = HC^\top (CC^\top)^{-1} y \tag{1}$$

is the conditional expectation (or the maximum likelihood estimate) of z given $Cd = y$.

However, it is possible to interpret this estimate of z without assuming randomness, as follows. 'Explain' the observed y as being generated by the d of least Euclidean norm that satisfies $y = Cd$. Denote this least squares d by d^*, and define the resulting estimate $\hat{z}$ of z by $\hat{z} = Hd^*$. It is elementary to verify that this yields again (1) as the formula for $\hat{z}$.

Which interpretation of (1) is to be preferred, the probabilistic one with its conditional mean/maximum likelihood interpretation, or the deterministic least squares one, is a matter of debate. Indeed, it has been a matter of debate at least since Gauss (1995) introduced least squares, or should one say *justified* Legendre's (1805) least squares, as a method of computing the most probable, maximum likelihood, outcome. It is not my purpose to re-open this debate, although I would like to state that I find, and have always found, simple least squares more satisfactory. It is more pragmatic, and it lays its strengths and weaknesses bare. All too often, probabilistic reasoning evokes a thick cloud of evasive claims concerning statistical knowledge about uncertainty. The uncertainty in models is very often due to such things as model approximation and simplification, neglected dynamics of sensors, unknown deterministic inputs, etc. It is hard to conceive situations in which precise stochastic knowledge about real uncertain disturbance signals can be justified as a description of reality.

Furthermore, in the case of continuous-time Kalman filtering, the deterministic least squares approach that I will present has also major pedagogical advantages. It totally avoids white noise, Brownian motion, stochastic calculus, and all that. That the deterministic approach is reasonable and convincing as a methodology, perhaps more so than the stochastic approach, may be to some extent a matter of opinion. However, the pedagogical advantages are beyond debate.

4. Filtering

In filtering problems, there are two time-signals involved: an *observed* signal, denoted by y, and a *to-be-estimated* signal, denoted by z. Both are assumed to be vector-valued, and, at first, we will take $\mathbb{R}_+$ as the time set on which these signals are defined. Later on, the case that the time set is $\mathbb{R}$ will also be discussed, and the development will make clear what happens when the time set is a finite interval.

Hence, let $y : \mathbb{R}_+ \to \mathbb{R}^y$ be the observed signal, and $z : \mathbb{R}_+ \to \mathbb{R}^z$ be the to-be-estimated signal. The basic problem is find a map $\mathscr{F} : y \mapsto \hat{z}$ such that $\hat{z} : \mathbb{R}_+ \to \mathbb{R}^z$ is a good

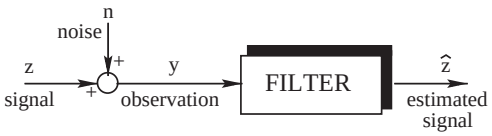

Fig. 1. Noise suppression.

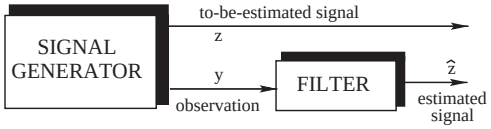

Fig. 2. Filter configuration.

estimate of z. The requirement that makes the filtering problem interesting and difficult is the constraint that the estimate $\hat{z}(T)$ at time T is allowed to depend only on the *past* of y, i.e., on the observed outcomes $y(t)$ for $0 \leqslant t \leqslant T$. In other words, the filter map $\mathscr{F}$ is required to be *non-anticipating*.

The map F is called a *filter*. In the earliest formulations, this problem was considered one of noise suppression, with $y = z + n$ and n unwanted '*noise*' (see Fig. 1). The problem then is to 'filter' the noise out of the observations.

Nowadays, the picture of Fig. 2 is more appropriate. Indeed, it is usually assumed that the to-be-estimated signal and the observed signal can be related in a more general way, and that both are outputs of a common signal generator.

The filtering problem as intuitively formulated above is of obvious relevance in a large variety of situations. However, in order to turn it into a mathematical question, we need to:

1. Model the relation between y and z mathematically.
2. Formulate an estimation principle.
3. Obtain an algorithm that computes $\hat{z}$ from y, i.e., an algorithm that implements the filter map $\mathscr{F}$.

In the stochastic approach to filtering, the relationship between z and y is specified by assuming that (z, y) is a stochastic process with known statistics, (i.e., it is assumed that the probability distributions of the relevant random variables are given). The estimation principle is then typically the requirement that $\hat{z}(T)$ must be the conditional expectation of $z(T)$ given $y(t)$ for $0 \leqslant t \leqslant T$. The Kalman filter formulas provide an effective, beautiful algorithm for computing $\hat{z}$ from y. The Kalman filter formulas are obtained by assuming that the stochastic model that yields (z, y) is given through a Gauss–Markov model, more specifically, through a linear system driven by disturbance inputs, and with these disturbances modeled as white noise processes (this model will be described more explicitly in Section 9).

The deterministic approach to filtering discussed in the present paper uses a similar model, but without the stochastic assumptions. Assume that the signals (z, y) are

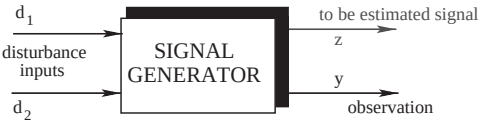

Fig. 3. Signal generation.

generated by the linear system (see Section 13 for an introduction to this model class)

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Gd_1, \quad y = Cx + d_2, \quad z = Hx. \tag{2}$$

Here $A \in \mathbb{R}^{n\times n}, G \in \mathbb{R}^{n\times d}, C \in \mathbb{R}^{y\times n}$, and $H \in \mathbb{R}^{z\times n}$ are fixed, known,[1] matrices that parameterize the system, and $d_1 : \mathbb{R}_+ \to \mathbb{R}^d$, $d_2 : \mathbb{R}_+ \to \mathbb{R}^y$, $x : \mathbb{R}_+ \to \mathbb{R}^n$, $y : \mathbb{R}_+ \to \mathbb{R}^y$, and $z : \mathbb{R}_+ \to \mathbb{R}^z$ are signals that are related through this linear system (2) (see Fig. 3).

The vector signal (d_1, d_2) should be interpreted as an (unobserved) disturbance input, which, together with the (unobserved) initial state $x(0) \in \mathbb{R}^n$ of (2), determines the observed signal y and the to-be-estimated signal z. More precisely, assume that $d_1 \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$ and $d_2 \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^y)$, then y and z are given in terms of $(d_1, d_2), x(0)$, and the system parameter matrices (A, G, C, H), by

$$y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Gd_1(\tau)\,\mathrm{d}\tau + d_2(t), \tag{3}$$

$$z(t) = He^{At}x(0) + \int_0^t He^{A(t-\tau)}Gd_1(\tau)\,\mathrm{d}\tau \tag{4}$$

for $t \geqslant 0$. It is easy to see from these expressions that $y \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^y)$ and $z \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^z)$.

The signal generation model hence assumes that there is a 'hidden' vector signal (d_1, d_2) and a 'hidden' initial state $x(0)$ which, through (3), generate the observed signal y and, through (4), the to-be-estimated signal z. We will return to the interpretation of the model in Section 10.4.

In the sequel, in order to distinguish between an arbitrary output y and the output that is observed, we will denote the output that has actually been observed by $\tilde{y}$. The problem is to find a non-anticipating filter map $\mathscr{F} : \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^y) \to \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^z)$, so that $\mathscr{F}(\tilde{y})(T)$ is a good estimate of $z(T)$.

[1] In filtering problems, the system model (in case, the matrices (A, G, C, H)) is assumed to be known. Of course, the problem of deducing the model from observations is also of much interest. In system theory, it is called *system identification*. It is a problem to which Manfred Deistler has made important contributions (Hannan and Deistler, 1988). For a treatise of identification algorithms from an applications oriented perspective, see Ljung (1987).

5. The least squares filtering principle

Assume that the output $\tilde{y}: \mathbb{R} \to \mathbb{R}^y$ has been observed. We want to obtain a filter that maps $\tilde{y}$ into the estimate $\hat{z}$.

What is a rational way of obtaining an estimate

$\hat{z}(T)$ *of* $z(T)$ *from* $\tilde{y}(t)$ *for* $0 \leqslant t \leqslant T$?

I have already described the stochastic approach: declare (d_1, d_2), and $x(0)$ to be random, and use conditional expectation. The deterministic approach put forward in this article is based on the following principle.

1. Among all the $d_1 \in \mathscr{L}_2([0,T], \mathbb{R}^d)$, $d_2 \in \mathscr{L}_2([0,T], \mathbb{R}^y)$, and $x(0) \in \mathbb{R}^n$ that '*explain*' the observed $\tilde{y} \in \mathscr{L}_2([0,T], \mathbb{R}^y)$, compute the one that minimizes the norm squared

$$\|x(0)\|_\Gamma^2 + \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}. \tag{5}$$

Here $\Gamma \in \mathbb{R}^{n\times n}$ is a given matrix that satisfies $\Gamma = \Gamma^\top \succ 0$. As usual, $\|\cdot\|_\Gamma$ denotes the norm on $\mathbb{R}^n$ defined by $\|a\|_\Gamma^2 := a^\top \Gamma a$. The above functional (5) is called the *uncertainty measure*. Alternative choices of the functional to be minimized will be discussed in Section 10.

With '*explain*', I mean that in this minimization, only those $(d_1, d_2), x(0)$ are considered, which, after substitution in (3), yield the observed signal $\tilde{y}$, i.e., such that

$$\tilde{y}(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Gd_1(\tau)\,d\tau + d_2(t) \tag{6}$$

for $0 \leqslant t \leqslant T$.

2. Denote the minimizing $(d_1, d_2), x(0)$, obtained in step 1, by $(d_1^*, d_2^*), x(0)^*$. Now, substitute $(d_1^*, d_2^*), x(0)^*$ in (4). Denote the resulting output by z^*. Hence

$$z^*(t) = He^{At}x(0)^* + \int_0^t He^{A(t-\tau)}Gd_1^*(\tau)\,d\tau$$

for $0 \leqslant t \leqslant T$.

3. Define the desired estimate of $z(T)$ by $\hat{z}(T) := z^*(T)$, with z^* obtained in step 2. Hence

$$\hat{z}(T) = He^{AT}x(0)^* + \int_0^T He^{A(T-\tau)}Gd_1^*(\tau)\,d\tau.$$

The principle underlying this procedure is reasonable and intuitively quite acceptable: among all $(d_1, d_2), x(0)$ that explain the observations, choose the one that has '*smallest uncertainty measure*', that is '*most likely*', where *'most likely'* should be interpreted as '*of least squares norm*', with the norm chosen as the square root of (5). Note that it is obvious from this construction that $\hat{z}(T)$ depends only on $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$. Hence the filter map $\mathscr{F}: \tilde{y} \mapsto \hat{z}$ is non-anticipating.

The construction of the optimal estimate $\hat{z}$ appears quite involved, since we need to compute $z^*(T)$ for all $T \in \mathbb{R}_+$. So, it appears as if, *at each time* T, we have to carry out, in step 1, what looks like a complicated dynamic optimization problem in order to compute $(d_1^*, d_2^*), x(0)^*$. The optimization problem is complicated indeed, but we will

see that it admits a very nice recursive solution, which makes it possible to carry out the computation of $\hat{z}$ in a very efficient way, *and for all T at once*!

6. Completion of the squares

The crucial lemmas that yield the solution of the minimization problem set up in the previous section use the Riccati differential equation. For the benefit of the uninitiated reader, and in order to make the paper self-contained, I have included the required know-how about the Riccati differential equation in Section 15.

Lemma 1. *Consider the following system of differential equations involving* $d_1 : [0,T] \to \mathbb{R}^d$, $d_2 : [0,T] \to \mathbb{R}^y$, $x : [0,T] \to \mathbb{R}^n$, $y : [0,T] \to \mathbb{R}^y$, $\hat{x} : [0,T] \to \mathbb{R}^n$, and $\Sigma : [0,T] \to \mathbb{R}^{n \times n}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Gd_1, \quad y = Cx + d_2,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + \Sigma C^\top (y - C\hat{x}),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma = GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma.$$

Then, assuming that $d_1 \in \mathscr{L}_2([0,T],\mathbb{R}^d)$, $d_2 \in \mathscr{L}_2([0,T],\mathbb{R}^y)$, *and that* $\Sigma(t) \in \mathbb{R}^{n \times n}$ *is symmetric and non-singular for* $0 \leqslant t \leqslant T$, *there holds*

$$\begin{aligned}
&\|x(0) - \hat{x}(0)\|^2_{\Sigma(0)^{-1}} + \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)} \\
&\quad = \|x(T) - \hat{x}(T)\|^2_{\Sigma(T)^{-1}} + \|d_1 - G^\top \Sigma^{-1}(x - \hat{x})\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} \\
&\qquad + \|y - C\hat{x}\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}.
\end{aligned}$$

Proof. Verify the following straightforward calculation

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}[(x - \hat{x})^\top \Sigma^{-1}(x - \hat{x})] \\
&\quad = 2(x - \hat{x})^\top \Sigma^{-1} \frac{\mathrm{d}}{\mathrm{d}t}(x - \hat{x}) + (x - \hat{x})^\top (\frac{\mathrm{d}}{\mathrm{d}t}\Sigma^{-1})(x - \hat{x}) \\
&\quad = 2(x - \hat{x})^\top \Sigma^{-1}[A(x - \hat{x}) + Gd_1 - \Sigma C^\top C(x - \hat{x}) - \Sigma C^\top d_2] \\
&\qquad + (x - \hat{x})^\top (C^\top C - A^\top \Sigma^{-1} - \Sigma^{-1}A - \Sigma^{-1}GG^\top \Sigma^{-1})(x - \hat{x}) \\
&\quad = -(x - \hat{x})^\top C^\top C(x - \hat{x}) + 2(x - \hat{x})^\top \Sigma^{-1} Gd_1 \\
&\qquad - 2(x - \hat{x})^\top C^\top d_2 - (x - \hat{x})^\top \Sigma^{-1} GG^\top \Sigma^{-1}(x - \hat{x}) \\
&\quad = \|d_1\|^2 + \|d_2\|^2 - \|d_1 - G^\top \Sigma^{-1}(x - \hat{x})\|^2 - \|y - C\hat{x}\|^2,
\end{aligned}$$

and integrate. □

Now, specialize the above lemma by specifying the initial value of the Riccati differential equation for Σ and of the differential equation for $\hat{x}$. For the initial value of Σ, use $\Sigma(0) = \Gamma^{-1}$. Since $\Gamma = \Gamma^\top \succ 0$, it follows from Theorem 7 that the Riccati differential equation with this initial condition has a unique solution on the interval $[0, T]$, and that this solution is symmetric and positive definite.

Lemma 2. *Let $\Sigma : [0, T] \to \mathbb{R}^{n\times n}$ be the (unique) solution to the Riccati differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma = GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma, \Sigma(0) = \Gamma^{-1}. \tag{7}$$

Then $\Sigma(t) = \Sigma(t)^\top \succ 0$ for $0 \leqslant t \leqslant T$.

Consider the system of differential equations involving $d_1 : [0, T] \to \mathbb{R}^d, d_2 : [0, T] \to \mathbb{R}^y, x : [0, T] \to \mathbb{R}^n$, and $y : [0, T] \to \mathbb{R}^y$

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Gd_1, \quad y = Cx + d_2.$$

Assume that $d_1 \in \mathscr{L}_2([0, T], \mathbb{R}^d)$, $d_2 \in \mathscr{L}_2([0, T], \mathbb{R}^y)$. Then $y \in \mathscr{L}_2([0, T], \mathbb{R}^y)$.

Define $\hat{x}$ in terms of Σ and y by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + \Sigma C^\top (y - C\hat{x}), \quad \hat{x}(0) = 0.$$

Then

$$\begin{aligned}\|x(0)\|^2_\Gamma + \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)} \\ = \|x(T) - \hat{x}(T)\|^2_{\Sigma(T)^{-1}} + \|d_1 - G^\top \Sigma^{-1}(x - \hat{x})\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} \\ + \|y - C\hat{x}\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}\end{aligned} \tag{8}$$

Proof. This follows immediately from the previous lemma. □

7. The least squares filter

The optimal filter is readily deduced from Lemma 2. Indeed, (8) shows that, whenever $d_1 \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$, $d_2 \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^y)$, and $x(0) \in \mathbb{R}^n$ lead to the observed signal $\tilde{y} \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^y)$, there holds

$$\|x(0)\|^2_\Gamma + \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)} \geqslant \|\tilde{y} - C\hat{x}\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}, \tag{9}$$

with $\hat{x}$ generated from $\tilde{y}$ by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + \Sigma C^\top (\tilde{y} - C\hat{x}), \quad \hat{x}(0) = 0, \tag{10}$$

and Σ defined by (7).

Observe, very importantly, that $\hat{x}$ is a function of $\tilde{y}$, but that it does not depend on the specific (d_1, d_2), and $x(0)$ that generated $\tilde{y}$. Hence the right hand side of inequality (9) depends on $\tilde{y}$ only. Therefore

$$\|x(0)\|^2_\Gamma + \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}$$

will be minimized if equality holds in (9). Lemma 2 shows that equality holds if and only if, among the (d_1, d_2), and $x(0)$ that generated $\tilde{y}$, we can choose one such that

1. $x(T) = \hat{x}(T)$, and
2. $d_1(t) = G^\top \Sigma(t)^{-1}(x(t) - \hat{x}(t))$ for $0 \leqslant t \leqslant T$.

Such a choice indeed exists! It is intuitively quite clear from 1 and 2 how to make this choice, and it will be formally derived in the proof of Theorem 3.

Note the following important consequence of the existence of such a choice. It implies in particular that the optimal $(d_1^*, d_2^*), x(0)^*$ yields for the state trajectory generated by $(d_1^*, x(0)^*)$, at time T, $x(T) = \hat{x}(T)$, and hence $\hat{z}(T) = H\hat{x}(T)$. Therefore, once we prove the mere existence of a choice (d_1, d_2), and $x(0)$ that meets the above conditions 1 and 2, the specific expressions of the optimal $(d_1^*, d_2^*), x(0)^*$ are not needed any further: (10) and $\hat{z} = H\hat{x}$ yield the optimal filter.

This leads to the main result of this paper.

Theorem 3 (Least squares filter). *Consider the least squares filtering problem defined by the linear system* (2) *with the uncertainty measure* (5). *Let* $\tilde{y} \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^y)$ *be an observed output. Let* $\Sigma : \mathbb{R}_+ \to \mathbb{R}^{n \times n}$ *be the* (*unique*) *solution of the Riccati differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma = GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma, \quad \Sigma(0) = \Gamma^{-1}. \tag{11}$$

Then the least squares filter is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + \Sigma C^\top(\tilde{y} - C\hat{x}), \quad \hat{x}(0) = 0, \quad \hat{z} = H\hat{x}, \tag{12}$$

viewed as map $\tilde{y} \in \mathscr{L}_2([0, \infty), \mathbb{R}^y) \mapsto \hat{z} \in \mathscr{L}_2([0, \infty), \mathbb{R}^z)$.

Proof. It will be shown that there exist unique $d_1 \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^d)$, $d_2 \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^y)$, and $x(0) \in \mathbb{R}^n$ such that the corresponding solution $x : [0, T] \to \mathbb{R}^n$ to $(\mathrm{d}/\mathrm{d}t)x = Ax + Gd_1$, $y = Cx + d_2$ yields (i) $y(t) = \tilde{y}(t)$ for $0 \leqslant t \leqslant T$, (ii) $x(T) = \hat{x}(T)$, and (iii) $d_1(t) = G^\top \Sigma(t)^{-1}(x(t) - \hat{x}(t))$ for $0 \leqslant t \leqslant T$.

This choice is obtained by first solving the Riccati differential equation (1) with initial condition $\Sigma(0) = \Gamma$ 'forwards' on $[0, T]$ (this yields $\Sigma(t)$ for $t \in [0, T]$), then the filter equation (12) with initial condition $\hat{x}(0) = 0$ 'forwards' on $[0, T]$ (this yields $\hat{x}(t)$ for $t \in [0, T]$), and finally solving

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x} = A\tilde{x} - GG^\top \Sigma^{-1}(\tilde{x} - \hat{x}), \quad \tilde{x}(T) = \hat{x}(T), \tag{13}$$

'backwards' on $[0, T]$ (this yields $\tilde{x}(t)$ for $t \in [0, T]$).

Now define the optimal (d_1, d_2), and $x(0)$ by

$$d_1^*(t) = G^\top \Sigma(t)^{-1}(\tilde{x}(t) - \hat{x}(t)), \quad \text{for } 0 \leqslant t \leqslant T,$$

$$d_2^*(t) = \tilde{y}(t) - C\tilde{x}(t), \quad \text{for } 0 \leqslant t \leqslant T, \text{ and}$$

$$x(0)^* = \tilde{x}(0).$$

It is straightforward to verify that

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Gd_1^*(t), \quad y = Cx + d_2^*(t), \quad x(0) = x(0)^*$$

yields $x(t) = \tilde{x}(t)$, for $0 \leqslant t \leqslant T$. Hence (i) $Cx(t) + d_2^*(t) = C\tilde{x}(t) + d_2^*(t) = \tilde{y}(t)$ for $0 \leqslant t \leqslant T$, (ii) $x(T) = \tilde{x}(T) = \hat{x}(T)$, and (iii) $d_1^*(t) = G^\top \Sigma(t)^{-1}(\tilde{x}(t) - \hat{x}(t)) = G^\top \Sigma(t)^{-1}(x(t) - \hat{x}(t))$ for $0 \leqslant t \leqslant T$. Optimality follows.

To see the uniqueness, note that $x(T) = \hat{x}(T)$ and $d_1(t) = G^\top \Sigma(t)^{-1}(x(t) - \hat{x}(t))$ for $0 \leqslant t \leqslant T$ are satisfied only if $x(t) = \tilde{x}(t)$, with $\tilde{x}$ governed by (13).

The above formulas show how to compute the inputs $d_1^* : [0, T] \to \mathbb{R}^d, d_2^* : [0, T] \to \mathbb{R}^y$, and the initial state $x(0)^*$ that optimally (in the least squares sense) explain the observations $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$. In principle, we should now compute $\hat{z} : \mathbb{R}_+ \to \mathbb{R}^z$ by solving the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Gd_1^*(t), \quad x(0) = x(0)^*$$

on $[0, T]$, computing $\hat{z}(T) = Hx(T)$, and *repeating this for each* $T \in \mathbb{R}_+$. *But there is no need to do that, since* d_1^* *and* $x(0)^*$ *were chosen so that there would hold* $x(T) = \hat{x}(T)$. Hence $\hat{z}(T) = H\hat{x}(T)$ for all $T \geqslant 0$.

The (somewhat surprising) conclusion that can drawn from the above is that there is no need to repeat the calculation for each T, and that (12) is the optimal filter. □

8. Infinite-time filtering

There are two ways to turn the filtering problem into an infinite-time problem: either by simply letting $t \to \infty$ in (12), or by assuming that y is observed on all of $\mathbb{R}$, and that the filter computes the optimal estimate $\hat{z}(T)$ on the basis of the observations $\tilde{y}(t)$ for all $t \leqslant T$.

The latter version, i.e., when the time-set is $\mathbb{R}$, is discussed first. In this case, it is necessary to assume some regularity on the system (2), for example, stability. Assume first that it is stable (see Section 13), i.e., that $A \in \mathbb{R}^{n \times n}$ is Hurwitz. This implies, in particular, that for any $d_1 \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^d), d_2 \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$, there exist unique $x \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$, $y \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$, and $z \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^z)$ such that (2) holds. This $x : \mathbb{R} \to \mathbb{R}^n$ is given by

$$x(t) = \int_{-\infty}^{t} \mathrm{e}^{A(t-\tau)} G d_1(\tau)\, \mathrm{d}\tau.$$

This yields

$$y(t) = \int_{-\infty}^{t} C\mathrm{e}^{A(t-\tau)} G d_1(\tau)\, \mathrm{d}\tau + d_2(t),$$

$$z(t) = \int_{-\infty}^{t} H\mathrm{e}^{A(t-\tau)} G d_1(\tau)\, \mathrm{d}\tau$$

for the corresponding $y \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$ and $z \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^z)$.

The least squares filtering problem may now be formulated as follows:

1. Let $\tilde{y}: \mathbb{R} \to \mathbb{R}^y$ be observed, and assume that $\tilde{y} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$.
2. Among all the $d_1 \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^d), d_2 \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$ that explain this $\tilde{y}$ in the sense that
$$\tilde{y}(t) = \int_{-\infty}^{t} C\mathrm{e}^{A(t-\tau)}Gd_1(\tau)\,\mathrm{d}\tau + d_2(t),$$
compute the one that minimizes the *uncertainty measure*, defined as the norm squared
$$\|d_1\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R}^y)}.$$
Denote this minimizing (d_1, d_2) by (d_1^*, d_2^*).
3. Define the estimate $\hat{z} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^z)$ by
$$\hat{z}(T) = \int_{-\infty}^{T} H\mathrm{e}^{A(T-\tau)}Gd_1^*(\tau)\,\mathrm{d}\tau.$$

This problem can be solved in an analogous way as the finite-time problem. The solution in this case uses the algebraic Riccati equation. The relevant facts about the Riccati equation are briefly reviewed in Section 16.

Theorem 4 (Infinite-time least squares filter). *Consider the least squares filtering problem defined by the linear system* (2) *with* $A \in \mathbb{R}^{n\times n}$ *Hurwitz, and the uncertainty measure* (21). *Let* $\tilde{y} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$ *be an observed output. Let* $\sum_\infty \in \mathbb{R}^{n\times n}$ *be the* (*unique*) *solution of the algebraic Riccati equation*

$$GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma = 0, \quad \Sigma = \Sigma^\top \geqslant 0. \tag{14}$$

Then the infinite-time least squares filter is given through the unique solution $\hat{x} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$ *of*

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + \sum_\infty C^\top(\tilde{y} - C\hat{x}), \quad \hat{z} = H\hat{x}. \tag{15}$$

From the theory of the algebraic Riccati equation (*see proposition* 8), *it follows that* $A - \sum_\infty C^\top C$ *is Hurwitz. Therefore* (15) *has a unique solution* $\hat{x} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$, *and so the filter is well-defined. In fact*,

$$\hat{z}(t) = \int_{-\infty}^{t} H\mathrm{e}^{(A-\sum_\infty C^\top C)(t-\tau)}\sum_\infty C^\top \tilde{y}(\tau)\,\mathrm{d}\tau.$$

Proof. Only an outline of the proof is given.

The result follows readily, analogously to the finite-time case, when $\sum_\infty \succ 0$. Unfortunately, it is only guaranteed that $\sum_\infty \geqslant 0$. The proof for the case $\sum_\infty \succ 0$ is given first, and later on, it will be indicated how it needs to be modified for the general case.

Note that $\hat{x} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$ and $\tilde{y} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$ imply, using (15), that $(\mathrm{d}/\mathrm{d}t)\hat{x} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$. Now, $\hat{x}, (\mathrm{d}/\mathrm{d}t)\hat{x} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$ implies $\hat{x}(t) \to 0$ as $t \to -\infty$. Now, repeat the

steps leading to Eq. (8), and obtain

$$\|d_1\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R}^y)} = \|x(T) - \hat{x}(T)\|^2_{\sum_\infty^{-1}}$$
$$+\|d_1 - G^\top \Sigma^{-1}(x - \hat{x})\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R}^d)} + \|y - C\hat{x}\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R}^y)}. \tag{16}$$

Now, verify that the proof used in the finite-time case goes through with the obvious changes.

In the general case, when $\sum_\infty \geqslant 0$, prove first that $(x - \hat{x})(t) \in \mathrm{im}(\sum_\infty)$ for all $t \in \mathbb{R}$, and deduce then the analogue of (16) with $\sum_\infty^{-1}$ replaced by $\sum_\infty^{\#}$, with $\sum_\infty^{\#}$ the pseudo-inverse of $\sum_\infty$. The details are omitted. □

The above formulation of the infinite-time filtering problem uses stability of the signal generator in an essential way. We may avoid this (unpleasant) assumption by assuming instead that $d_1, d_2, x, y, \tilde{y}$, and z are in $\mathscr{L}_2^-$ and of compact support, and that the system is detectable and stabilizable (see Section 14).

In the second way of approaching the infinite-time problem, by considering the problem on $[0, T]$ and letting $T \to \infty$, stability is not needed, and detectability and stabilizability are sufficient. Assume that the pair of matrices (A, C) is detectable, and that the pair of matrices (A, G) is stabilizable. It follows from the theory of the Riccati equation (see Section 16) that, in this case, there exists a unique solution to the algebraic Riccati equation

$$GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma = 0$$

that is symmetric and non-negative definite: $\Sigma = \Sigma^\top \geqslant 0$. Denote this solution by $\sum_\infty$. Moreover, for any initial condition $\Sigma_0 = \Sigma_0^\top \geqslant 0$, the solution of the Riccati differential equation (see Section 16)

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma = GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma, \quad \Sigma(0) = \Sigma_0$$

converges to $\sum_\infty$: $\Sigma(t) \to \sum_\infty$ as $t \to \infty$.

It follows that when (A, C) is detectable and (A, G) is stabilizable, then the filter derived in Theorem 3 approaches, as $t \to \infty$, the filter

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} + \sum_\infty C^\top(\tilde{y} - C\hat{x}), \quad \hat{z} = H\hat{x}. \tag{17}$$

This filter has very good properties. In particular, $A - C^\top C \sum_\infty$ is Hurwitz, and so, the filter is stable. This implies that the estimate $\hat{z}(t)$ in Eq. (15) becomes independent of $x(0)$ for $t \to \infty$. This asymptotic independence can be viewed as a form of '*merging of opinions*' (of the initial condition $\Sigma(0) = \Gamma$ on the Riccati equation and the initial condition $\hat{x}(0)$ of the filter—see Remark 10.1. The dynamics of the estimation error $e = z - \hat{z}$ may be obtained by subtracting (17) from (2). This yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}_x = \left(A - C^\top C \sum_\infty\right)\mathrm{e}_x + Gd_1 - \sum_\infty C^\top d_2, \quad \mathrm{e} = H\mathrm{e}_x.$$

This shows, for example, that if d_1 and d_2 have compact support, then $\hat{z}(t) - z(t) \to 0$ for $t \to \infty$. This is interesting in the case that A is not Hurwitz, since then z and $\tilde{y}$

may be unbounded, but nevertheless the filter output $\hat{z}$ tracks z asymptotically without error: while $z(t), \hat{z}(t) \to \infty$ for $t \to \infty$, $z(t) - \hat{z}(t) \to 0$ for $t \to \infty$.

9. Stochastic interpretation

It is well-known, of course, that the filter derived in Theorem 3 is also obtained, with exactly the same formulas, under the following stochastic assumptions.

1. $x(0)$ is a gaussian random vector with zero mean and covariance matrix Γ.
2. (d_1, d_2) is a zero-mean white noise process with identity covariance, and independent of $x(0)$. The system equations are then usually written, for the sake of mathematical rigor, as a stochastic differential equation

$$\mathrm{d}x = Ax\,\mathrm{d}t + G\,\mathrm{d}w_1, \quad \mathrm{d}f = Cx\,\mathrm{d}t + \mathrm{d}w_2, \quad z = Hx,$$

   with (w_1, w_2) a Wiener process, and it is assumed that it is f (instead of $y = (\mathrm{d}/\mathrm{d}t)f$) that is observed.
3. $\hat{z}(T)$ is the conditional expectation, equivalently, the maximum likelihood estimate, of $z(T)$ given the observations $y(t)$ for $0 \leqslant t \leqslant T$.

The least squares approach has important advantages above the stochastic approach. It is a very rational and reasonable principle in itself, and avoids modeling the uncertainty probabilistically. In applications, this uncertainty is often due to effects as quantization, saturation and other neglected (non-linear) features of the plant and the sensors, inputs with an unknown and/or unmeasured origin, etc. In such situations, the probabilistic interpretation is heuristic, at best, but not an attempt to describe reality.

Pedagogically, the least squares approach also has a great deal to be said for. It allows one to dispense with the difficult aspects of the mathematical etiquette that occurs when one has to consider stochastic differential equations. It allows to aim much more directly at the Kalman filter as a signal processor.

It is of interest to interpret the functional (5), or, perhaps more appropriately,

$$\mathrm{e}^{-(\|x(0)\|_I^2 + \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^{d})} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^{y})})}$$

as a *belief*, or a *likelihood function* that expresses numerically the degree of confidence in the joint occurrence of $x(0)$ and $(d_1, d_2)(t)$ for $0 \leqslant t \leqslant T$. Unfortunately, to the best of my knowledge, the classical axiomatization of belief functions (Paris, 1994) does not cover this interpretation. Note that, informally, also the stochastic white noise interpretation leads to a likelihood in which realizations of (d_1, d_2) and $x(0)$ for which the norm (5) is large, are viewed as less likely than realizations for which the norm is small. However, since realizations of white noise have—with probability one—infinite $\mathscr{L}_2$ norm on any interval, this interpretation is destined to remain an informal one. In fact, our derivation of the deterministic Kalman filter can be considered a mathematical proof that, in the stochastic case, the maximum likelihood estimate of $x(T)$ given the observations $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$ can be computed by deterministic least squares. It is an interesting question to examine if this principle extends for more general Itô equations.

10. Remarks

There are number of directions in which the results of this paper can be extended.

10.1.

Consider the filtering problem in which, instead of (5), the uncertainty functional

$$\|x(0)-a\|_{\Gamma}^2+\|d_1\|_{\mathscr{L}_2([0,T],\mathbb{R}^d)}^2+\|d_2\|_{\mathscr{L}_2([0,T],\mathbb{R}^y)}^2$$

is minimized, with $a\in\mathbb{R}^n$ a fixed vector, and, as before, $\Gamma\in\mathbb{R}^{n\times n}, \Gamma=\Gamma^\top\succ 0$. The optimal filter is then identical to the one of (12), but with initial condition $\hat{x}(0)=a$. This corresponds completely to the usual situation considered in the Kalman filter.

10.2.

The formulas are easily adapted, as in the stochastic Kalman filter, to the case in which the disturbance inputs d_1 and d_2 are coupled. The system equations then take the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x=Ax+Gd,\quad y=Cx+Dd,\quad z=Hx,$$

and $\|d_1\|_{\mathscr{L}_2([0,T],\mathbb{R}^d)}^2+\|d_2\|_{\mathscr{L}_2([0,T],\mathbb{R}^y)}^2$ in (5) is replaced by $\|d\|_{\mathscr{L}_2([0,T],\mathbb{R}^d)}^2$. The filter remains identical (assuming that D has full row rank), but the relevant Riccati differential equation for Σ becomes more involved.

10.3.

Often, there is, in addition to a measured output, y, also a measured input, u, leading to the system equations

$$\frac{\mathrm{d}}{\mathrm{d}t}x=Ax+Bu+Gd,\quad y=Cx+Dd,\quad z=Hx.$$

In this case, it suffices to add Bu to the right hand side of the filter equations (12) in order to obtain the least squares filter.

10.4.

Let us reflect once again about the general problem formulation of filtering, and obtain a point of view that has all the above situations as special cases.

Assume that we have two (vector) signals, $w:[0,T]\to\mathbb{R}^w$ and $z:[0,T]\to\mathbb{R}^z$. We think of w as a signal, that we will, perhaps imperfectly, observe, and through which we wish to estimate the unobserved signal z. The first thing to do, is model the (ideal) relation between w and z. A reasonable model, in the context of linear systems, is

$$\frac{\mathrm{d}}{\mathrm{d}t}x=Ax+Bu+Gd,\quad y=Cx+Dd,\quad w=(u,y),\quad z=Hx. \tag{18}$$

In this model we have partitioned w into a free input u and a bound output y. Such a partition is, in a precise sense, always possible (see Polderman and Willems, 1998). But, more importantly, we have introduced another free input d. We view this input as a *latent* input, which together with the initial state $x(0)$, yields a well-defined relation between $w=(u,y)$ and z. This latent input is a *passe-partout* term that expresses model inadequacies.

Now assume that we measure the signal w. Let $\tilde{w}$ be the vector signal that is actually observed. In order to take into consideration things as modeling approximations, saturation, discretizations and quantizations, nonlinear effects, sensor dynamics and inaccuracies, etc., it is reasonable not to jump to the conclusion that $\tilde{w}$ must satisfy the model equations, but to assume that there is another signal, $\hat{w}$, that satisfies the model equations (in other words for which there is an $(d,x(0))$ which together with $\hat{w}$ satisfies the model equations) and for which the observed $\tilde{w}$ is but an imperfect manifestation.

This brings about two elements through which the actual observations are explained:

(i) the *misfit* $\tilde{w}-\hat{w}$, measured, say, as $\|\tilde{w}-\hat{w}\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^w)}$, and
(ii) the *latency* $(d,x(0))$, measured, say, as $\|x(0)-a\|^2_\Gamma+\|d\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)}$.

Excessive reliance on either of these, the misfit or the latency, in order to explain the measurements is, in principle, undesirable. It is therefore reasonable to determine the optimal $(d^*,x(0)^*)$ and w^*, as the one that minimizes the (weighted) sum of the misfit and the latency, yielding the *uncertainty measure*

$$\|\tilde{w}-\hat{w}\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^w)}+\|x(0)-a\|^2_\Gamma+\|d\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)}.$$

The filter algorithm obtained in Section 7 is readily extended to this situation, and yields a non-anticipating filtering algorithm that obtains $\hat{z}$ from $\tilde{w}$.

For the deterministic Kalman filter discussed in Section 5, the input u is absent, and it is most reasonable to consider $(d_1,x(0))$ as the latency, measured as $\|x(0)\|^2_\Gamma+\|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)}$, and $\tilde{y}-Cx$ as the misfit, measured as $\|\tilde{y}-Cx\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}$, whence assuming in effect that Cx is a signal that is generated by $(d_1,x(0))$ and that is measured through y with misfit $d_2=y-Cx$.

I consider the above as the most reasonable interpretation of the Kalman filter that I have come across.[2]

Note that it is not unreasonable to interpret this situation in terms of *fuzzy logic*. Eqs. (18) define a *behavior*: all trajectories (d,x,Hx,Cx) that are declared possible, that are compatible with these equations. We have two fuzzy membership functions. An *internal fuzzy membership* function which tells how likely an element of the behavior will occur, and an *external fuzzy membership* function, which tells how close a trajectory comes to belonging to the behavior. The latency (or, better, $e^{-(\|x(0)\|^2_\Gamma+\|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)})}$) defines the internal fuzzy membership, while the misfit (or, better, $e^{-\|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}}$) defines the external fuzzy membership. Our filter looks for the maximum of the total

[2] Note the close relation of this interpretation with *errors-in-variables* filtering, another area to which Manfred Deistler has made important contributions.

fuzzy membership function (the product of the internal and the external memberships) that is compatible with the observations.

10.5.

It is well-known that in the stochastic case, $\Sigma(T)$ can be interpreted as a measure for the uncertainty of the estimate $\hat{x}(T)$, in the sense that $\Sigma(T)$ equals the expected value of $(x(T)-\hat{x}(T))(x(T)-\hat{x}(T))^\top$.

A similar interpretation holds in the deterministic case. Assume that, as in Section 7, we set out to estimate $z(T)$ from the observations $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$. As we have seen, this leads to estimating $x(T)$ through the optimal $x(0)$ and d_1. We can do this in two batches: process first the observations for $0 \leqslant t \leqslant t_1$, and subsequently for $t_1 \leqslant t \leqslant T$. It is in combining these two batches that $\Sigma(t_1)$ becomes important. Indeed, the correct way of proceeding is to estimate first $\hat{x}(t_1)$ using the algorithm of Theorem 3 by minimizing the partial uncertainty measure $\|x(0)\|_I^2 + \|d_1\|^2_{\mathscr{L}_2([0,t_1],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,t_1],\mathbb{R}^y)}$, and then estimate $\hat{x}(T)$ using the algorithm of Theorem 3 by minimizing (taking into consideration the modification explained in Remark 10.1) the partial uncertainty measure $\|x(t_1)-\hat{x}(t_1)\|^2_{\Sigma(t_1)^{-1}} + \|d_1\|^2_{\mathscr{L}_2([t_1,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([t_1,T],\mathbb{R}^y)}$. Thus $\|x(t)-\hat{x}(t)\|^2_{\Sigma(t)^{-1}}$ has the interpretation of the equivalent accumulated uncertainty (more exactly, the equivalent latency) at time t.

10.6.

The theory leading to the finite-time filter (12) goes through, at the expense of only more complicated notation, for time-varying systems.

10.7.

The theory leading to the finite-time least squares filter goes through, with suitable adaptation of the notation, for discrete time systems. In this case, the least squares filter allows a genuine interpretation as a maximum likelihood, conditional mean, stochastic estimator. However, it remains awkward to interpret the infinite-time filter of Theorem 4 stochastically.

10.8.

In *filtering*, it is assumed, as we have seen, that $z(T)$ is estimated on the basis of the observations $y(t)$ for $0 \leqslant t \leqslant T'$ with $T' = T$. The estimation problem is also of interest when $T' \neq T$. When $T' < T$, the problem is called *prediction*, when $T' > T$, is called *smoothing*. The least squares filtering approach is readily extended to cover these situations as well. We should then minimize (5) with $T = T_f$, under the constraint that $x(0), d_1, d_2$ explains the observations.

10.9.

Note that in Theorem 4 we assumed $\tilde{y} \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^y)$. Of course, the resulting filter can be shown to be optimal for other function spaces as well, for example, for (almost) periodic signals. More pragmatically, it is reasonable to use a filter that works well for $\mathscr{L}_2$ signals, also for other signals for which its dynamics are well-defined.

10.10.

We are presently investigating how the theory of quadratic differential forms (Willems and Trentelman, 1998) can be used to cover more general uncertainty measures than (5). This extension is analogous to assuming (d_1, d_2) to be colored noise in traditional stochastic filtering.

10.11.

I believe that the deterministic least squares interpretation of filtering will have important (conceptual and algorithmic) advantages when considering multi-dimensional systems, e.g., images, or distributed phenomena governed by partial differential equations. In these applications, the stochastic interpretation of uncertainty becomes often even more tenuous.

11. Examples

We now discuss a couple of quick examples in order to illustrate the action of the deterministic Kalman filter intuitively. Note first that if, instead of the uncertainty (5), the weighted sum $\|x(0)\|_I^2 + \rho^2 \|d_1\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T],\mathbb{R}^y)}$ with weighting $\rho > 0$ is used, then it suffices to replace $GG^\top$ in the Riccati equation by $GG^\top/\rho^2$. This also pertains to the infinite-time case.

11.1. An integrator

Consider the scalar model $(\mathrm{d}/\mathrm{d}t)x = d_1$, $y = x + d_2$, and the uncertainty measure

$$\rho^2 \|d_1\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R})} + \|d_2\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R})}.$$

The corresponding algebraic Riccati equation is $\Sigma^2 - 1/\rho^2 = 0$, whence $\Sigma_\infty = 1/\rho$. The filter becomes $(\mathrm{d}/\mathrm{d}t)\hat{x} = \frac{1}{\rho}(\tilde{y} - \hat{x})$: exponential weighting with time constant ρ, whence $\hat{x}(T) = \int_0^\infty \frac{1}{\rho} \mathrm{e}^{-t/\rho} \tilde{y}(T-t)\,\mathrm{d}t$.

11.2. An interpretation

We can interpret this example as a model for the price y of a share of stock. This price is modeled as the sum of the intrinsic value x and a term d_2 depending on the market mood. The intrinsic value fluctuates as a consequence of change d_1 of the fundamentals. The weighting factor ρ reflects the relative volatility of the change of the

fundamentals and the market mood. Assuming, for example, that the mood fluctuates in a day as the fundamentals do in a month, leads to $\rho = 30$ days. The optimal estimate of the intrinsic value from the share trading price is then the exponential weighting of the share price with a time constant of 30 days. Note that our interpretation as a deterministic Kalman filter yields this exponential smoothing, without need to enter into a tenuous interpretation of x, y, d_1, d_2 as stochastic processes.

11.3. A double integrator

Consider the state model $(\mathrm{d}/\mathrm{d}t)x_1 = x_2, (\mathrm{d}/\mathrm{d}t)x_2 = d$, with d an (un-measured, latent) input. Let $\tilde{y}$ be the measurement, and define the uncertainty measure as

$$\rho^2 \|d\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R})} + \|\tilde{y} - x_1\|^2_{\mathscr{L}_2((-\infty,T],\mathbb{R})},$$

with $\rho > 0$ a weighting factor. The algebraic Riccati equation in

$$\Sigma = \begin{bmatrix} \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_3 \end{bmatrix}$$

yields $2\sigma_2 = \sigma_1^2, \sigma_2^2 = 1/\rho^2, \sigma_3 = \sigma_1\sigma_2$, resulting in

$$\Sigma_\infty = \begin{bmatrix} \sqrt{2/\rho} & 1/\rho \\ 1/\rho & \sqrt{2}/\rho\sqrt{\rho} \end{bmatrix}.$$

The infinite-time filter becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x}_1 = \hat{x}_2 + \sqrt{2/\rho}(\tilde{y} - \hat{x}_1), \qquad \frac{\mathrm{d}}{\mathrm{d}t}\hat{x}_2 = 1/\rho(\tilde{y} - \hat{x}_1).$$

The transfer functions $\tilde{y} \mapsto \hat{x}_2$ and $\tilde{y} \mapsto \hat{x}_1$ are equal to

$$\frac{s}{\rho s^2 + \sqrt{2\rho}s + 1} \quad \text{and} \quad \frac{\sqrt{2\rho}\, s + 1}{\rho\, s^2 + \sqrt{2\rho}s + 1},$$

respectively. The filter $\tilde{y} \mapsto \hat{x}_2$ acts as a differentiator at low frequencies, but quenches high frequencies, with the turn-around point at frequency $\sim 1/2\pi\sqrt{\rho}$.

11.4. An interpretation

We can interpret this example as the estimation of the velocity of a moving vehicle (for example, a truck) by measuring its position (for example, from a satellite). The vehicle accelerates and decelerates in unpredictable ways (due to gear shifting, braking, hills and other terrain irregularities, etc.). Assume also that the measurements can only recognize the position with a limited accuracy and are quantized in time and space. The aim is to estimate the velocity. A somewhat surprising, but very useful, achievement of the Kalman filter is that it provides an algorithm that makes it possible to estimate the velocity, not by differentiating (generally a non-robust operation, and obviously unsuitable for the case at hand), but by using the system equations instead.

Figs. 4–11 show the result of a simulation. The motion is assumed to be one dimensional (the vehicle drives along a straight road). Fig. 4 shows the acceleration of the

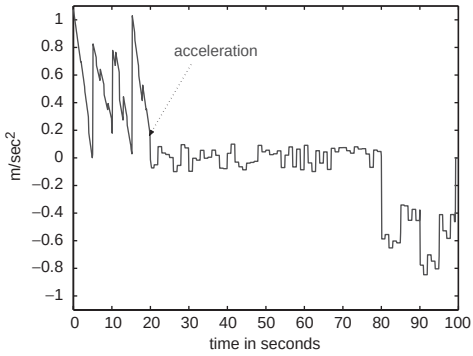

Fig. 4. Acceleration.

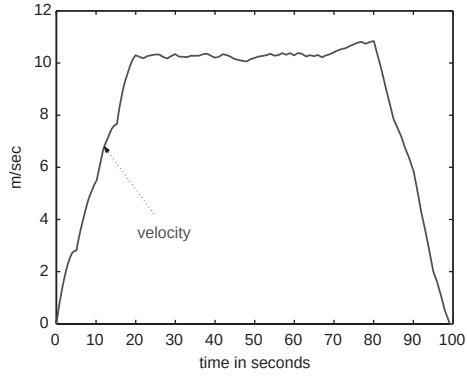

Fig. 5. Velocity.

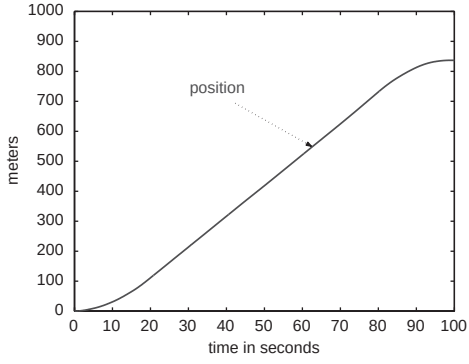

Fig. 6. Position.

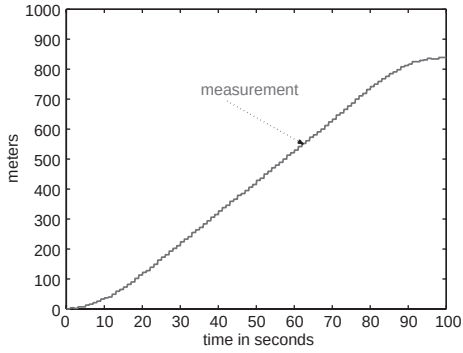

Fig. 7. Measurement.

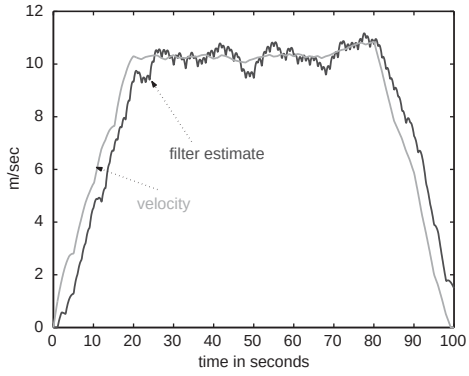

Fig. 8. Filter estimate.

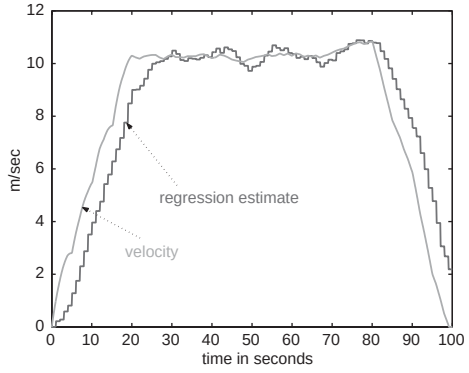

Fig. 9. Regression estimate

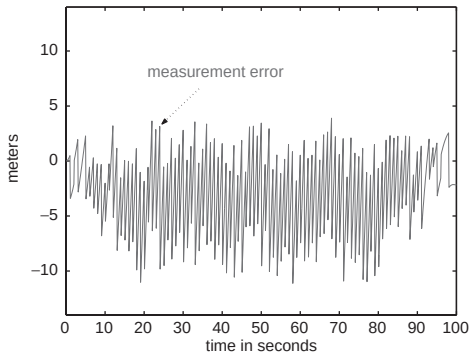

Fig. 10. Measurement errors.

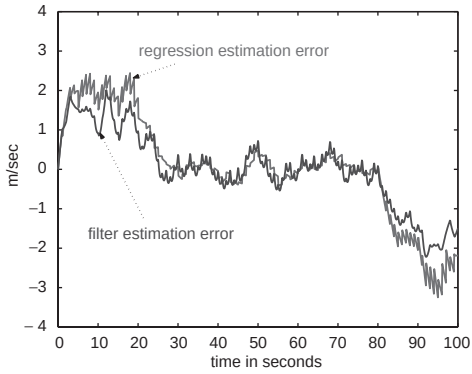

Fig. 11. Estimation errors.

vehicle, Fig. 5 its velocity, and Fig. 6 its position. The measurements of the position, shown in Fig. 7, are assumed to be taken once a second, and rounded up to 1 m after being corrupted by an additive noise term with a 5 m range (in order to express the size of the moving object). The resulting velocity estimates, obtained by the algorithm of Section 11.3 with $\rho = 5$, are shown in Fig. 8. In order to evaluate the performance of the filter, we also computed the velocity estimate obtained by fitting at each instant the optimal least squares regression line to the 10 most recent measurements (the number 10 being chosen to obtain good performance: much less than 10 leads to very noisy regression estimates, and much more than 10 leads to a very sluggish estimator). The resulting estimates are shown in Fig. 9. The measurement error is shown in Fig. 10, and the estimation errors of both the deterministic Kalman filter and the regression are shown in Fig. 11. Observe that the Kalman filter performs somewhat better than the estimator obtained by regression, especially during acceleration or deceleration. However, from the computational point of view, the Kalman filter requires integration of merely a second order system, compared to the regression estimator which effectively comes down to using a 10th order moving average filter.

Fig. 12. Plant.

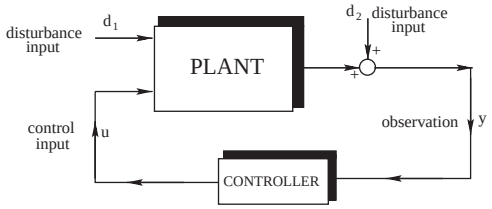

Fig. 13. Feedback control.

12. Least squares control

The least squares principle that I have put forward for filtering, can also be used for control. This is explained in the present section. Consider the *plant*

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Bu + Gd_1, \qquad y = Cx + d_2. \tag{19}$$

Here $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times u}, G \in \mathbb{R}^{n\times d}$, and $C \in \mathbb{R}^{y\times n}$ are fixed, known matrices that parameterize the system, and $u : \mathbb{R} \to \mathbb{R}^u, d_1 : \mathbb{R}_+ \to \mathbb{R}^d$, $d_2 : \mathbb{R}_+ \to \mathbb{R}^y$, $x : \mathbb{R}_+ \to \mathbb{R}^n$, and $y : \mathbb{R}_+ \to \mathbb{R}^y$ are signals that are related through the linear system equations (19) (see Fig. 12).

The vector signal u is a to-be-chosen control input, while (d_1, d_2) should, as in the filtering problem, be interpreted as an (unobserved) disturbance input. Together with the (unobserved) initial state $x(0) \in \mathbb{R}^n$ of (19), these inputs determine the observed signal y.

In *feedback control* (see Fig. 13), the control input u should be chosen on the basis of observations y, so as achieve one or other control objective, for example, steer the state (or another output) along a desirable path. Denote the *controller* by $\mathscr{N} : y \mapsto u$. The requirement that makes the feedback control problem interesting and difficult is the (natural) constraint that the control input value $u(T) = \mathscr{N}(y)(T)$ at time T is allowed to depend only on the *past* of y, i.e., on the observed outcomes $y(t)$ for $0 \leqslant t \leqslant T$. In other words, the controller map $\mathscr{N}$ is required to be *non-anticipating*.

Formally, a *feedback controller* is defined as a map[3]

$$\mathscr{N} : \mathscr{L}_2([0, T_f], \mathbb{R}^y) \to \mathscr{L}_2([0, T_f], \mathbb{R}^u)$$

[3] This choice of *admissible* feedback controllers is adequate for least squares control. However, in other applications, one may need to consider a more general class, allowing, for example, also differential operators.

that is *non-anticipating*. This means that $y_1, y_2 \in \mathscr{L}_2([0,T_f],\mathbb{R}^y)$, $T \in [0,T_f]$, and[4] $a_1(t) = a_2(t)$ for $0 \leqslant t \leqslant T$ imply $\mathscr{N}(a_1)(t) = \mathscr{N}(a_2)(t)$ for $0 \leqslant t \leqslant T$.

In *linear-quadratic* (LQ) *control*, the controller performance is expressed by means of a *cost-functional*, given by an integral of a quadratic expression in u, x, and the terminal state $x(T_f)$. The cost will be taken to be

$$\int_0^{T_f} [u(t)^\top R u(t) + x(t)^\top L x(t)]\,\mathrm{d}t + x(T_f)^\top Q x(T_f). \tag{20}$$

Here $T_f > 0$ is the *control horizon*, and $R \in \mathbb{R}^{u\times u}, L \in \mathbb{R}^{n\times n}, Q \in \mathbb{R}^{n\times n}$ are the *cost matrices*. Assume $R = R^\top \succ 0, L = L^\top$, and $Q = Q^\top$.

The problem is to select the feedback controller that minimizes the cost (20). However, the value of the cost does not depend solely on $\mathscr{N}$, but also on (d_1,d_2) and $x(0)$, and so, we meet the common situation of having to optimize a functional that depends both on the decision variable ($\mathscr{N}$) and uncertainty ($x(0),d_1,d_2$). As in the filtering problem, these are only partially and indirectly known through the observations. In general, of course, there will not be a feedback controller that minimizes the cost for all (d_1,d_2) and $x(0)$. The classical way out of this dilemma is to assume that (d_1,d_2) and $x(0)$ are stochastic, and to choose $\mathscr{N}$ so as to minimize the expected value of the cost. For the case at hand, the problem of finding the optimal stochastic feedback controller is very nicely solved in what is called the *linear-quadratic-gaussian* (LQG) *problem* (see, for example Athans, 1971; Kwakernaak and Sivan, 1972; Wonham, 1968; Willems, 1978).

In this section, I will show that this optimal control problem allows a perfectly reasonable and rational, deterministic least squares formulation.

The basic methodological question to be answered is

When do we want to call the feedback controller $\mathscr{N}^*$ optimal?

Optimality will be based on the following principle.

1. Fix a time $T \in [0,T_f]$ in the control horizon interval. Among all $d_1 \in \mathscr{L}_2([0,T_f],\mathbb{R}^d)$, $d_2 \in \mathscr{L}_2([0,T_f],\mathbb{R}^y)$, and $x(0) \in \mathbb{R}^n$ that '*explain*' the already observed $\tilde{y} \in \mathscr{L}_2([0,T),\mathbb{R}^y)$, compute the one that *at time* T minimizes the *uncertainty measure*

$$\|x(0)\|_\Gamma^2 + \|d_1\|^2_{\mathscr{L}_2([0,T_f],\mathbb{R}^d)} + \|d_2\|^2_{\mathscr{L}_2([0,T_f],\mathbb{R}^y)}. \tag{21}$$

Here $\Gamma \in \mathbb{R}^{n\times n}$ is a given matrix that satisfies $\Gamma = \Gamma^\top \succ 0$.

With '*explain*', I mean, as in the filtering case, that only those $(d_1,d_2),x(0)$ are considered, which, after substitution in (19), yield the observed signal $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$, i.e., such that

$$\tilde{y}(t) = C\mathrm{e}^{At}x(0) + \int_0^t C\mathrm{e}^{A(t-\tau)}B\mathscr{N}^*(\tilde{y})(\tau)\,\mathrm{d}\tau + \int_0^t C\mathrm{e}^{A(t-\tau)}Gd_1(\tau)\,\mathrm{d}\tau + d_2(t),$$

for $0 \leqslant t \leqslant T$.

[4] Here, and elsewhere, the equality $y_1|_{[0,T]} = y_2|_{[0,T]}$ and $\mathscr{N}(y_1)|_{[0,T]} = \mathscr{N}(y_2)|_{[0,T]}$ should be understood in the $\mathscr{L}_2$-sense.

2. Denote the minimizing $(d_1, d_2), x(0)$, obtained in step 1, by $(d_1^*, d_2^*), x(0)^*$. Note that we now have chosen (d_1^*, d_2^*) optimally on the whole interval $[0, T_f]$, since the integrals in (21) extend over the whole interval $[0, T_f]$. On $[0, T]$, their prediction is obtained analogously to the filtering case, while on $[T, T_f]$, they will be taken to be equal to zero. Now, solve the following optimal control problem. Compute the control input that minimizes the cost-functional (20), subject to

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Bu + Gd_1^*, \quad x(0) = x(0)^*$$

over all $u \in \mathscr{L}_2([0, T_f], \mathbb{R}^u)$ such that $u(t) = \mathscr{N}^*(\tilde{y})(t)$ for $0 \leqslant t \leqslant T$. Note that in this minimization, it is assumed that the disturbance $d_1 = d_1^*$ and the initial state $x(0) = x(0)^*$ are known, equal to those that were declared 'most likely' in step 1.

This minimization problem is what is called an '*open-loop*' *control problem*. We are basically simply minimizing over all possible control input signals $u \in \mathscr{L}_2([0, T_f], \mathbb{R}^u)$, or, more to the point, over all control-to-go input signals $u|_{[T, T_f]} \in \mathscr{L}_2([T, T_f], \mathbb{R}^u)$.

3. Denote the optimal control input obtained in step 2 by u^*. Of course, u^* depends on d_1^* and $x(0)^*$, which in turn depend on $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$. It will be shown that u^* is continuous from the right at T. Define $\bar{u}(T) = \lim_{t \downarrow T}(u^*(t))$. Of course, $\bar{u}(T)$ also depends on d_1^* and $x(0)^*$, which in turn depend on $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$. The optimality condition for $\mathscr{N}^*$ is

$$\bar{u} = \mathscr{N}^*(\tilde{y}) \text{ for all } \tilde{y} \in \mathscr{L}_2([0, T_f], \mathbb{R}^y)$$

The principle underlying this procedure is reasonable and intuitively quite acceptable: among all $(d_1, d_2), x(0)$ that explain the observations $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$, choose the one that has '*smallest uncertainty measure*', that is '*most likely*', where '*most likely*' should be interpreted as '*of least squares norm*', with the norm chosen as the square root of (21). Then compute the control-to-go that minimizes the cost (20), under the assumption that the disturbance d_1 and $x(0)$ are given by their most likely value, and that the control $\tilde{u}(t) = \mathscr{N}^*(\tilde{y})(t)$ for $0 \leqslant t \leqslant T$ (these control values have already been decided upon at time T). This yields $\bar{u}(T)$ as the 'best' control value that should be used at time T. If, for all $\hat{y}$ and all T, the resulting decision $\bar{u}(T)$ corresponds to $\mathscr{N}^*(\tilde{y})(T)$, as $\mathscr{N}^*$ would have suggested, then $\mathscr{N}^*$ is declared to be optimal.

This construction of $\bar{u}$ appears quite involved, since it requires that we compute it for all $T \in [0, T_f]$. So, it appears as if, *at each time* T, we have to carry out, both in step 1 and in step 2, what look like complicated dynamic optimization problems. These optimization problems are complicated indeed, but we will see that they admit very nice recursive solutions, which makes it possible to carry out the computation in a very efficient way.

The following theorem gives the solution of the least squares control problem.

Theorem 5 (Least squares controller). *Consider the least squares control problem defined by the plant* (19), *the uncertainty measure* (21), *and the cost-functional* (20). *Let* $\Sigma : \mathbb{R}_+ \to \mathbb{R}^{n \times n}$ *be the* (*unique*) *solution of the filtering Riccati differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma = GG^\top + A\Sigma + \Sigma A^\top - \Sigma C^\top C\Sigma, \quad \Sigma(0) = \Gamma^{-1}. \tag{22}$$

Assume further that the control Riccati differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}K = -L - A^\top K - KA + KBR^{-1}B^\top K, \quad K(T_f) = Q \tag{23}$$

has a solution $K : [0, T_f] \to \mathbb{R}^{n\times n}$ (*this solution is then unique and symmetric*: $K(t) = K^\top(t)$, *and existence is guaranteed if, for example,* $L = L^\top \geqslant 0$ *and* $Q = Q^\top \geqslant 0$)).

Then the controller defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = A\hat{x} - BR^{-1}B^\top K\hat{x} + \Sigma C^\top(\tilde{y} - C\hat{x}), \hat{x}(0) = 0, \tilde{u} = -R^{-1}B^\top K\hat{x}, \tag{24}$$

viewed as map $\tilde{y} \in \mathscr{L}_2([0, T_f], \mathbb{R}^y) \mapsto \tilde{u} \in \mathscr{L}_2([0, T_f], \mathbb{R}^u)$, *is optimal.*

Proof. Only an outline of the proof is given.

Step 1: The first step of the proof computes the least squares filter with a controller in the loop. Let $\mathscr{N}$ be a feedback controller, $\tilde{y} \in \mathscr{L}_2(\mathbb{R}, \mathbb{R}^y)$ an observed output of the system

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + B\mathscr{N}(y) + Gd_1, y = Cx + d_2,$$

and $0 \leqslant T \leqslant T_f$. Consider now the problem of explaining the observations $\tilde{y}(t)$ for $0 \leqslant t \leqslant T$ optimally in the sense of minimizing (21) over all $d_1, d_2, x(0)$ such that

$$\tilde{y}(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}B\mathscr{N}(\tilde{y})(\tau)\,\mathrm{d}\tau + \int_0^t Ce^{A(t-\tau)}Gd_1(\tau)\,\mathrm{d}\tau + d_2(t)$$

for $0 \leqslant t \leqslant T$. Repeating the proof of (3) yields for the optimum

$$d_1^*(t) = \begin{cases} G^\top \Sigma(t)^{-1}(\tilde{x}(t) - \hat{\hat{x}}(t)), & \text{if } 0 \leqslant t \leqslant T, \\ 0 & \text{otherwise,} \end{cases}$$

$$d_2^*(t) = \begin{cases} \tilde{y}(t) - C\hat{\hat{x}}(t), & \text{for } 0 \leqslant t \leqslant T, \\ 0 & \text{otherwise,} \end{cases}$$

$$x(0)^* = \tilde{x}(0),$$

where $\hat{\hat{x}}$ and $\tilde{x}$ are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\hat{x}} = A\hat{\hat{x}} + \mathscr{N}(\tilde{y}) + \Sigma C^\top(\tilde{y} - C\hat{\hat{x}}), \quad \hat{\hat{x}}(0) = 0,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x} = A\tilde{x} + \mathscr{N}(\tilde{y}) + GG^\top \Sigma^{-1}(\tilde{x} - \hat{\hat{x}}), \quad \tilde{x}(T) = \hat{\hat{x}}(T).$$

Step 2: The second step of the proof consists of identifying the optimal control-to-go. We need the following lemma.

Lemma 6. *Consider the following system of differential equations involving* $u : [0, T_f] \to \mathbb{R}^u, d_1 : [0, T_f] \to \mathbb{R}^d$, $x : [0, T_f] \to \mathbb{R}^n$, $s : [0, T_f] \to \mathbb{R}^n$, *and* $K : [0, T_f] \to \mathbb{R}^{n \times n}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Bu + Gd_1,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}s = -As + KBR^{-1}B^\top s - KGd_1,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}K = -L - A^\top K - KA + KBR^{-1}B^\top K.$$

Then there holds

$$\frac{\mathrm{d}}{\mathrm{d}t}(x^\top Kx + 2x^\top s) + \|u\|_R^2 + \|x\|_L^2$$

$$= \|u + R^{-1}B^\top(Kx + s)\|_R^2 - \|B^\top s\|_{R^{-1}}^2 + 2s^\top Gd_1.$$

Proof. After substitution of $(\mathrm{d}/\mathrm{d}t)x, (\mathrm{d}/\mathrm{d}t)s$, and $(\mathrm{d}/\mathrm{d}t)K$, the proof is a straightforward calculation, along the lines of the proof of Lemma 1. □

It follows from the above lemma that, with $s(T_f) = 0$, there holds

$$\int_0^{T_f} [u(t)^\top Ru(t) + x(t)^\top Lx(t)]\,\mathrm{d}t + x(T_f)^\top Qx(T_f)$$

$$= \int_0^{T_f} |u + R^{-1}B^\top(Kx + s)|_R^2\,\mathrm{d}t$$

$$+ x(0)^\top K(0)x(0) + 2x(0)^\top s(0) - \int_0^{T_f} |B^\top s|_{R^{-1}}^2\,\mathrm{d}t + \int_0^{T_f} 2s^\top Gd_1\,\mathrm{d}t.$$

Now, consider the cost (20), with $x(0)$ and $d_1 \in \mathscr{L}_2([0, T_f], \mathbb{R}^d$ given, and $s(T_f) = 0$. It follows that on the right hand side of the above expression, only the first term

$$\int_0^{T_f} |u + R^{-1}B^\top(Kx + s)|_R^2\,\mathrm{d}t$$

depends on the control $u \in \mathscr{L}_2(\mathbb{R}, \mathbb{R}^u)$.

This allows to solve the open-loop control problem of minimizing the cost-functional (20), subject to $(\mathrm{d}/\mathrm{d}t)x = Ax + Bu + Gd_1$, with d_1 and $x(0)$ given, over all $u \in \mathscr{L}_2([0, T_f], \mathbb{R}^u)$ such that $u(t) = \tilde{u}(t)$ for $0 \leqslant t \leqslant T$. It immediately follows that the optimal u^* is given by $u^*(t) = -R^{-1}B^\top(K\bar{x}(t) + s(t))$, for $T \leqslant t \leqslant T_f$, with

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{x} = A\bar{x} - BR^{-1}B^\top(K\bar{x} + s), \quad \bar{x}(T) = \tilde{x}(T)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x} = A\tilde{x} + B\tilde{u} + Gd_1, \quad \tilde{x}(0) = x(0).$$

Step 3: The above two steps provide the necessary ingredients for the proof of the theorem. Let $\mathscr{N}$ be a feedback controller, leading to an observed output $\tilde{y}$ and

$\tilde{u} = \mathscr{N}(\tilde{y})$. Let $T \in [0, T_f]$. Apply step 2 with $d_1 = d_1^*$, and $x(0) = x(0)^*$, with $d_1^*, x(0)^*$ obtained in step 1. Then, since $d_1^*(t) = 0$ for $T \leqslant t \leqslant T_f$ and $s(T_f) = 0$, the corresponding $s(t) = 0$ for $T \leqslant t \leqslant T_f$. Hence the optimal control-to-go is given by $u^*(t) = -R^{-1}B^\top K\bar{x}(t)$, for $T \leqslant t \leqslant T_f$, with

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{x} = A\bar{x} - BR^{-1}B^\top K\bar{x}, \quad \bar{x}(T) = \tilde{x}(T)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x} = A\tilde{x} + B\tilde{u} + Gd_1^*, \quad \tilde{x}(0) = x(0)^*.$$

Clearly u^* is continuous from the right at T, and $\lim_{t \downarrow T}(u^*(t)) = -R^{-1}B^\top K\tilde{x}(T)$. However, because of the choice of d_1^* and $x(0)^*$, there holds $\tilde{x}(T) = \hat{\tilde{x}}(T)$, with $\hat{\tilde{x}}$ given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\tilde{x}} = A\hat{\tilde{x}} + \mathscr{N}(\tilde{y}) + \Sigma C^\top(\tilde{y} - C\hat{\tilde{x}}), \quad \hat{\tilde{x}}(0) = 0.$$

Hence $\mathscr{N}$ is optimal if and only if it satisfies $\mathscr{N}(\tilde{y}) = -R^{-1}B^\top K\hat{\tilde{x}}$. It follows that the controller given in the statement of the theorem is indeed the optimal controller. □

Recapitulating, the optimal control law $\mathscr{N}^*$ is *incrementally* constructed as follows. Say that at time T, $\tilde{y}(t)$ has been observed for $0 \leqslant t < T$. From $\mathscr{N}^*$ (or by directly observing the control input, after all, we generated it ourselves), we can compute the corresponding control input $\tilde{u}(t)$ for $0 \leqslant t < T$. Estimate the $x(0)^*$ and $d_1^*(t), d_2^*(t)$ for $0 \leqslant t \leqslant T_f$ that optimally explain $\tilde{y}(t)$ for $0 \leqslant t < T$. Determine the optimal control-to-go, $\bar{u}(t)$ for $T \leqslant t \leqslant T_f$, by minimizing (20), with $x^*(T)$ and $d_1^*(t)$ obtained from the optimal estimation. Compute $\bar{u}(T)$. Then $u^*(T) = \bar{u}(T)$ yields the *incremental* continuation of the optimal control law.

The optimal control law that we have constructed satisfies a *separation* and a *certainty equivalence principle*, just as in the stochastic case, but albeit in a much more forced, more imposed, and hence mathematically less deep way. That the control law (24) satisfies a separation principle is obvious from its expression, which involves estimating $x(T)$, regardless of the control objective, and then using this estimate with control gains that are independent of the disturbance input matrix (G) or the observed output matrix (C) of the plant. Whence the filtering task may be regarded as separated from the control task.

That the control law (24) satisfies a certainty equivalence principle may be seen as follows. Let $y = x$. Assume that $\tilde{x}$ is observed. Then there is no need to estimate $x(0)$ and d_1, the optimal estimate will, of course, yield $x^*(T) = \tilde{x}(T)$ and $d_1^*(t) = 0$ for $T \leqslant t \leqslant T$. The optimal control, computed following Theorem 5, then becomes $\tilde{u}(T) = -R^{-1}B^\top K\tilde{x}(T)$. The optimal control law, for the relevant observations $y = Cx + d_1$, is given by $\tilde{u} = -R^{-1}B^\top K\hat{x}$. In other words, when the state x is not completely observed, it is replaced in the control law by its optimal estimate. The optimal control law acts 'equivalently' as if it this estimate were 'certain'. The value of the optimal control that is actually used is equal to the optimal least squares estimate of the optimal control that would be used if the whole state were observed.

13. Linear systems

Consider the linear time-invariant differential system

$$\frac{\mathrm{d}}{\mathrm{d}t}x = Ax + Bu, \qquad y = Cx + Du. \tag{25}$$

Here $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times u}, C \in \mathbb{R}^{y\times n}$, and $D \in \mathbb{R}^{y\times u}$ are fixed matrices that parameterize the system, $u : \mathbb{R} \to \mathbb{R}^u$ is the *input* trajectory, $x : \mathbb{R} \to \mathbb{R}^n$ is the *state* trajectory, while $y : \mathbb{R} \to \mathbb{R}^y$ is the *output* trajectory. In (25), u is considered an exogenous (vector) signal imposed on the system by its environment, while y is an endogenous (vector) signal, through which the system can in turn interact with its environment. The (vector) variable x is an intermediate variable introduced in order to specify the relation between u and y in a convenient way; x is called the state, because it succinctly expresses the *memory* of the system.

For any given input trajectory $u \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^u)$ and initial state $x(0) \in \mathbb{R}^n$, (25) defines a unique state trajectory $x \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ and output trajectory $y \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^y)$ given by

$$x(t) = \mathrm{e}^{At}x(0) + \int_0^t \mathrm{e}^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau,$$

$$y(t) = C\mathrm{e}^{At}x(0) + \int_0^t C\mathrm{e}^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau + Du(t).$$

The system (25) is said to be *stable* if $u = 0$ implies $x(t) \to 0$ for $t \to \infty$, equivalently, if $\mathrm{e}^{At} \to 0$ for $t \to \infty$. It is well-known that this is the case if and only if A is Hurwitz. A matrix $M \in \mathbb{R}^{n\times n}$ is said to be a *Hurwitz* matrix if all its eigenvalues have negative real part.

A stable system has the property that to each $u \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^u)$, there corresponds a *unique* state trajectory $x \in \mathscr{L}_2^-(\mathbb{R}, \mathbb{R}^n)$. It is given by

$$x(t) = \int_{-\infty}^t \mathrm{e}^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau.$$

The other state trajectories corresponding to this u, but with initial state $x(0)$ different from the one given by the above formula, are not square integrable on $(-\infty, 0]$.

14. Controllability and observability

In this section, the notions of controllability and observability, classical notions from the modern theory for dynamical systems, and their refinements, stabilizability and detectability, are reviewed. They are important in infinite-time filtering.

The system (25) is said to be *controllable* if it can be steered by the input from any state to any other state. Formally, if for any two states $x_0, x_1 \in \mathbb{R}^n$, there exists a $T > 0$, a continuous input $u : [0, T] \to \mathbb{R}^u$, and a continuously differentiable state trajectory $x : [0, T] \to \mathbb{R}^n$, such that $x(0) = x_0, x(T) = x_1$, and $((\mathrm{d}/\mathrm{d}t)x)(t) = Ax(t) + Bu(t)$ for all $t \in [0, T]$. Controllability can be expressed very succinctly in terms of the matrices A

and B. Indeed, it is easy to show (see, for example, Brockett, 1970, p. 80) that (25) is controllable if and only if the $n \times nu$ matrix $[B\ AB\ A^2B \cdots A^{n-1}B]$ has rank n.

The system (25) is said to be *stabilizable* if the system can be steered by the input from any state asymptotically to zero. Formally, if for any state $x_0 \in \mathbb{R}^n$, there exists a continuous input $u : \mathbb{R}_+ \to \mathbb{R}^u$, and a continuously differentiable state trajectory $x : [0, T] \to \mathbb{R}^n$, such that $x(0) = x_0$, $((\mathrm{d}/\mathrm{d}t)x)(t) = Ax(t) + Bu(t)$ for all $t \in [0, T]$, and $x(t) \to 0$ for $t \to \infty$. It is a bit more difficult to express stabilizability succinctly in terms of A and B. Perhaps the easiest characterization is that stabilizability of (25) is equivalent to the existence of a matrix $F \in \mathbb{R}^{u \times n}$ such that $A + BF$ is Hurwitz.

The system (25) is said to be *observable* if the state can be deduced from the input and output trajectories. In order to state this formally, define the *behavior* of (25) as $\mathscr{B} := \{u, x, y) \in \mathscr{L}_2^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^u \times \mathbb{R}^n \times \mathbb{R}^y) | (\mathrm{d}/\mathrm{d}t)x = Ax + Bu, y = Cx + Du$. Then (25) is said to be *observable* if $(u, x_1, y) \in \mathscr{B}$ and $(u, x_2, y) \in \mathscr{B}$ imply $x_1 = x_2$. Observability can be expressed very succinctly in terms of the matrices A and C. Indeed, it is easy to show (see, e.g., Brockett, 1970, p. 91) that (25) is observable if and only if the $n \times ny$ matrix $[C^\top A^\top C^\top (A^\top)^2 C^\top \cdots (A^\top)^{n-1} C^\top]$ has rank n.

The system (25) is said to be *detectable* if, asymptotically, the state can deduced from the input and output trajectories. Formally, (25) is said to be *detectable* if $(u, x_1, y) \in \mathscr{B}$ and $(u, x_2, y) \in \mathscr{B}$ imply $x_1(t) - x_2(t) \to 0$ as $t \to \infty$. It is a bit more difficult to express detectability succinctly in terms of A and C. Perhaps the easiest characterization is that detectability of (25) is equivalent to the existence of a matrix $L \in \mathbb{R}^{n \times y}$ such that $A + LC$ is Hurwitz.

Controllability & stabilizability, and observability & detectability have become such pervasive properties that are encountered in many contexts, that nowadays a *pair of matrices*(A, B), $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times u}$ is called *controllable* (*stabilizable*) if the corresponding system (25) is controllable (stabilizable). Similarly, a *pair of matrices* (A, C), $A \in \mathbb{R}^{n \times n}, C \in \mathbb{R}^{y \times n}$ is called *observable* (*detectable*) if the corresponding system (25) is observable (detectable).

15. The Riccati differential equation

The Riccati equation is a fixpoint in control theory. It is part of the algorithms that lead to the solution of the main problems in the field: the Kalman filter (Kalman, 1960a), linear-quadratic (LQ) (Kalman, 1960b) and linear-quadratic-gaussian (LQG) control (Wonham, 1968), the theory of dissipative systems (together with LMI's) (Willems, 1972), and $\mathscr{H}_\infty$-control. The remarkable solutions to the LQG and the $\mathscr{H}_\infty$-problems in terms of two Riccati equations (Doyle et al., 1989), are generally perceived as the most important results in control theory in the 60's and the 90's.

The Riccati equation is a non-linear differential equation that seems rather impenetrable when one first encounters it. It has as its unknown an $n \times n$ matrix, which I will denote by K, in honor of R.E. Kalman, who was instrumental in introducing this equation in control and filtering. What makes the Riccati equation difficult to grasp is the presence of (at least) 3 other $n \times n$ matrices, which, in view of the range of applications, one wants to leave in as general parameter matrices. These matrices are denoted as $Q, F, P \in \mathbb{R}^{n \times n}$.

Consider the nonlinear (quadratic) differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}K = Q + FK + KF^\top + KPK. \tag{26}$$

This should be viewed as a differential equation in the unknown matrix-valued function $K : \mathbb{R} \to \mathbb{R}^{n\times n}$, with $Q, F, P \in \mathbb{R}^{n\times n}$ fixed parameter matrices. This differential equation is called a *Riccati differential equation*, after Jacopo Riccati [1696–1754], who, around 1724, first studied this equation in the case $n = 1$.

Note that the map $K \in \mathbb{R}^{n\times n} \mapsto Q + FK + KF^\top + KPK \in \mathbb{R}^{n\times n}$ is a quadratic map. As such, the right hand side of (26) does not satisfy a global Lipschitz condition, and hence it is not guaranteed by the standard theory of differential equations that (26), with a specified initial value $K(0) \in \mathbb{R}^{n\times n}$, has a solution on all of $\mathbb{R}$. However, since this quadratic map is differentiable, there does exist a unique solution on a sufficiently small interval $[-\varepsilon, \varepsilon]$. Actually, it is easy to see that already in the case $n = 1$, with, for example, $Q = 1, F = 0, P = 1$, and $K(0) = 0$, there does not exist a solution on all of $\mathbb{R}_+$.

However, it is possible to give rather general conditions on Q, F, P and $K(0)$ so that a solution exists on at least the half line $\mathbb{R}_+$. The following result (by no means the best of its type) is adequate for the purposes of this paper.

Proposition 7. *Assume that* $Q, F, P, K_0 \in \mathbb{R}^{n\times n}$ *satisfy* (i) $Q = Q^\top \geqslant 0$, (ii) $P = P^\top \leqslant 0$, (iii) $K_0 = K_0^\top \geqslant 0$. *Then the Riccati differential equation* (26) *with initial condition* $K(0) = K_0$ *has a solution for* $t \geqslant 0$. *This solution is unique, symmetric, and non-negative definite*: $K(t) = K(t)^\top \geqslant 0$ *for all* $t \in \mathbb{R}_+$. *If* $K_0 = K_0^\top \succ 0$, *then, in fact, this solution is symmetric and positive definite*: $K(t) = K(t)^\top \succ 0$ *for all* $t \in \mathbb{R}_+$.

The proof is an exercise in the theory of differential equations (see, for example, Brockett, 1970, p. 165).

It turns out that for the filtering Riccati equation used in Theorems 3 and 5, existence of solutions is guaranteed by the above proposition. However, for the control Riccati differential equation used in Theorem 5, existence is left as an unresolved condition that needs to be established independently. Of course, application of the above proposition gives $L = L^\top \geqslant 0, Q = Q^\top \geqslant 0$ as a sufficient condition for existence.

16. The algebraic Riccati equation

When considering control or filtering problems over an infinite horizon, there is a nonlinear algebraic equation that takes over the role of the Riccati differential equation. This algebraic equation is

$$Q + FK + KF^\top + KPK = 0. \tag{27}$$

This equation should again be viewed as an equation in the unknown matrix $K \in \mathbb{R}^{n\times n}$, with $Q, F, P \in \mathbb{R}^{n\times n}$ fixed parameter matrices. This equation is called the *algebraic Riccati equation*. This name, while being a natural consequence of calling (26) the *Riccati differential equation*, may seem a bit strange. Indeed, (27) is a system of

quadratic equations, which, in the case $n = 1$, reduces to the quadratic equation that is studied in detail in high school algebra. This special case immediately shows that there may not exist real solutions, and, if a real solution exists, it is likely that it is not unique. Interestingly, the sufficiency conditions for existence involve controllability and observability considerations, more precisely, stabilizability and detectability.

Proposition 8. *Assume that* $Q, F, P \in \mathbb{R}^{n\times n}$ *satisfy* (i) $Q = Q^\top \geqslant 0$, (ii) $P = P^\top \leqslant 0$, (iii) (F, Q) *stabilizable, and* (iv) (F, P) *detectable. Then there exists a solution* $K \in \mathbb{R}^{n\times n}$ *to the algebraic Riccati equation* (27). *This solution is* (*in general*) *not unique, but there is a unique solution that is symmetric and non-negative definite*: $K = K^\top \geqslant 0$. *In other words, there is a unique solution to*

$$Q + FK + KF^\top + KPK = 0, K = K^\top \geqslant 0. \tag{28}$$

Moreover, this solution is such that $F + KP \in \mathbb{R}^{n\times n}$ *is Hurwitz.*

Note that the solutions of (27) are the equilibria of (26). As such, it is not surprising that the asymptotic behavior of the Riccati differential equation is very much related to the solutions of the algebraic Riccati equation. The following result to this effect is used in the theory of infinite-time filtering.

Proposition 9. *Assume that* $Q, F, P \in \mathbb{R}^{n\times n}$ *satisfy* (i) $Q = Q^\top \geqslant 0$, (ii) $P = P^\top \leqslant 0$, (iii) (F, Q) *stabilizable, and* (iv) (F, P) *detectable. Let* $K_\infty \in \mathbb{R}^{n\times n}$ *be the unique solution to the algebraic Riccati equation* (28). *Let* $K_0 \in \mathbb{R}^{n\times n}$ *satisfy* $K_0 = K_0^\top \geqslant 0$. *Then the Riccati differential equation* (27) *with initial condition* $K(0) = K_0$ *has a unique solution* $K : \mathbb{R}_+ \to \mathbb{R}^{n\times n}$, *and* $K(t) \to K_\infty$ *for* $t \to \infty$. *In other words, the set of non-negative definite symmetric real* $n \times n$ *matrices is included in the domain of attraction of the equilibrium* K_∞.

Acknowledgements

The author would like to thank Ivan Markovsky for some very useful discussions and comments. This research is supported by the Belgian Federal Government under the DWTC program Interuniversity Attraction Poles, Phase V, 2002–2006, Dynamical Systems and Control: Computation, Identification and Modelling.

References

Athans, M. (Ed.), 1971. Special issue on the linear-quadratic-gaussian estimation and control problem. IEEE Transactions on Automatic Control 16(6), 527–869.

Başar, T. (Ed.), 2001. Control Theory, Twenty-Five Seminal Papers (1932–1981), IEEE Press, New York.

Brockett, R.W., 1970. Finite Dimensional Linear Systems. Wiley, New York.

Doyle, J.C., Glover, K., Khargonekar, P., Francis, B.A., 1989. State space solutions to standard $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control problem. IEEE Transactions on Automatic Control 34, 831–847.

Fleming, W.H., 1997. Deterministic nonlinear filtering. Ann. Scuola Normale Superiore Pisa, Cl. Sci. Fis. Mat. 25, 435–454.

Gauss, C.F., 1995. Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, Dieterich, Göttingen, 1823. Recent translation by G.W. Stewart, Theory of the Combination of Observations Least Subject to Errors, Classics in Applied Mathematics, SIAM, Philadelphia, PA.

Hannan, E.J., Deistler, M., 1988. The Statistical Theory of Linear Systems. Wiley, New York.

Hijab, O., 1980. Minimum Energy Estimation. Ph.D. Dissertation, Department of Mathematics, University of California, Berkeley.

Kailath, T. (Ed.), 1977. Linear Least-Squares Estimation. Dowden, Hutchinson and Ross, Stroudburg.

Kalman, R.E., 1960a. A new approach to linear filtering and prediction problems. Transactions of the ASME, Journal of Basic Engineering 82D, 35–45.

Kalman, R.E., 1960b. Contributions to the theory of optimal control. Boletin de la Sociedad Matematica Mexicana 5, 102–119.

Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction theory. Transactions of the ASME, Journal of Basic Engineering 83D, 95–108.

Kolmogoroff, A., 1939. Sur l'interpolation et extrapolation des suites stationaires. Comptes Rendus de l'Académie des Sciences 208, 2043–2045.

Kwakernaak, H., Sivan, R., 1972. Linear Optimal Control Systems. Wiley, New York.

Legendre, A.M., 1805. Nouvelles Méthodes pour le Détermination de Orbites des Comètes. Courcier, Paris.

Ljung, L., 1987. System Identification: Theory for the User. Prentice-Hall, Englewood Cliffs, NJ.

McEneaney, W.M., 1998. Robust/$\mathscr{H}_\infty$ filtering of nonlinear systems. Systems Control Letters 33, 315–325.

Mortenson, R.E., 1968. Maximum likelihood recursive nonlinear filtering. Journal of Optimization Theory and Applications 2, 386–394.

Paris, J.B., 1994. The Uncertain Reasoner's Companion, A Mathematical Perspective. Cambridge University Press, Cambridge.

Polderman, J.W., Willems, J.C., 1998. Introduction to Mathematical Systems Theory: A Behavioral Approach. Springer, New York.

Sontag, E.D., 1990. Mathematical Control Theory, Deterministic Finite Dimensional Systems. Springer, Berlin.

Swerling, P., 1971. Modern state estimation methods from the viewpoint of the method of least squares. IEEE Transactions on Automatic Control 16 (6), 707–719.

Wiener, N., 1949. Extrapolation, Interpolation, and Smoothing of Stationary Time Series. MIT Press, Cambridge, MA.

Willems, J.C., 1972. Dissipative dynamical systems—Part I: General theory, Part II: Linear systems with quadratic supply rates. Archive for Rational Mechanics and Analysis 45, 321–351 and 352–393.

Willems, J.C., 1978. Recursive filtering. Statistica Neerlandica 32, 1–39.

Willems, J.C., Trentelman, H.L., 1998. On quadratic differential forms. SIAM Journal on Control and Optimization 36, 1702–1749.

Wonham, W.M., 1968. On the separation theorem of stochastic control. SIAM Journal on Control 6, 312–326.