We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the
TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us? Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

10

Nonlinear Prediction, Filtering and Smoothing

10.1 Introduction

The Kalman filter is widely used for linear estimation problems where its behaviour is well-understood. Under prescribed conditions, the estimated states are unbiased and stability is guaranteed. Many real-world problems are nonlinear which requires amendments to linear solutions. If the nonlinear models can be expressed in a state-space setting then the Kalman filter may find utility by applying linearisations at each time step. Linearising means finding tangents to the curves of interest about the current estimates, so that the standard filter recursions can be employed in tandem to produce predictions for the next step. This approach is known as extended Kalman filtering – see [1] – [5].

Extended Kalman filters (EKFs) revert to optimal Kalman filters when the problems become linear. Thus, EKFs can yield approximate minimum-variance estimates. However, there are no accompanying performance guarantees and they fall into the try-at-your-own-risk category. Indeed, Anderson and Moore [3] caution that the EKF "can be satisfactory on occasions". A number of compounding factors can cause performance degradation. The approximate linearisations may be crude and are carried out about estimated states (as opposed to true states). Observability problems occur when the variables do not map onto each other, giving rise to discontinuities within estimated state trajectories. Singularities within functions can result in non-positive solutions to the design Riccati equations and lead to instabilities.

The discussion includes suggestions for performance improvement and is organised as follows. The next section begins with Taylor series expansions, which are prerequisites for linearisation. First, second and third-order EKFs are then derived. EKFs tend be prone to instability and a way of enforcing stability is to masquerade the design Riccati equation by a faux version. This faux algebraic Riccati equation technique [6] – [10] is presented in Section 10.3. In Section 10.4, the higher order terms discarded by an EKF are treated as uncertainties. It is shown that a robust EKF arises by solving a scaled H_∞ problem in lieu of one possessing uncertainties. Nonlinear smoother procedures can be designed similarly. The use of fixed-lag and Rauch-Tung-Striebel smoothers may be preferable from a complexity perspective. However, the approximate minimum-variance and robust smoothers, which are presented in Section 10.5, revert to optimal solutions when the nonlinearities and uncertainties diminish. Another way of guaranteeing stability is to by imposing constraints and one such approach is discussed in Section 10.6.

"It is the mark of an instructed mind to rest satisfied with the degree of precision to which the nature of the subject admits and not to seek exactness when only an approximation of the truth is possible." *Aristotle*

10.2 Extended Kalman Filtering

10.2.1 Taylor Series Expansion

A nonlinear function $a_k(x): \mathbb{R}^n \to \mathbb{R}$ having n continuous derivatives may be expanded as a Taylor series about a point x_0

$$\begin{aligned} a_k(x) &= a_k(x_0) + \frac{1}{1!}(x - x_0)^T \nabla a_k(x_0) \\ &\quad + \frac{1}{2!}(x - x_0)^T \nabla^T \nabla a_k(x_0)(x - x_0) \\ &\quad + \frac{1}{3!}(x - x_0)^T \nabla^T \nabla (x - x_0) \nabla a_k(x_0)(x - x_0) \\ &\quad + \frac{1}{4!}(x - x_0)^T \nabla^T \nabla (x - x_0) \nabla (x - x_0) \nabla a_k(x_0)(x - x_0) + \ldots, \end{aligned} \tag{1}$$

where $\nabla a_k = \begin{bmatrix} \frac{\partial a_k}{\partial x_1} & \cdots & \frac{\partial a_k}{\partial x_n} \end{bmatrix}$ is known as the gradient of $a_k(.)$ and

$$\nabla^T \nabla a_k = \begin{bmatrix} \frac{\partial^2 a_k}{\partial x_1^2} & \frac{\partial^2 a_k}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 a_k}{\partial x_1 \partial x_n} \\ \frac{\partial^2 a_k}{\partial x_2 \partial x_1} & \frac{\partial^2 a_k}{\partial x_2^2} & \cdots & \frac{\partial^2 a_k}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 a_k}{\partial x_n \partial x_1} & \frac{\partial^2 a_k}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 a_k}{\partial x_n^2} \end{bmatrix}$$

is called a Hessian matrix.

10.2.2 Nonlinear Signal Models

Consider nonlinear systems having state-space representations of the form

$$x_{k+1} = a_k(x_k) + b_k(x_k) w_k, \tag{2}$$

$$y_k = c_k(x_k), \tag{3}$$

where $a_k(.)$, $b_k(.)$ and $c_k(.)$ are continuous differentiable functions. For a scalar function, $a_k(x): \mathbb{R} \to \mathbb{R}$, its Taylor series about $x = x_0$ may be written as

---

"In the real world, nothing happens at the right place at the right time. It is the job of journalists and historians to correct that." *Samuel Langhorne Clemens aka. Mark Twain*

$$a_k(x) = a_k(x_0) + (x - x_0)\left.\frac{\partial a_k}{\partial x}\right|_{x=x_0} + \frac{1}{2}(x - x_0)^2 \left.\frac{\partial^2 a_k}{\partial x^2}\right|_{x=x_0} + \frac{1}{6}(x - x_0)^3 \left.\frac{\partial^3 a_k}{\partial x^3}\right|_{x=x_0} + \cdots + \frac{1}{n!}(x - x_0)^n \left.\frac{\partial^n a_k}{\partial x^n}\right|_{x=x_0} . \tag{4}$$

Similarly, Taylor series for $b_k(x): \mathbb{R} \to \mathbb{R}$ and $c_k(x): \mathbb{R} \to \mathbb{R}$ about $x = x_0$ are

$$b_k(x) = b_k(x_0) + (x - x_0)\left.\frac{\partial b_k}{\partial x}\right|_{x=x_0} + \frac{1}{2}(x - x_0)^2 \left.\frac{\partial^2 b_k}{\partial x^2}\right|_{x=x_0} + \frac{1}{6}(x - x_0)^3 \left.\frac{\partial^3 b_k}{\partial x^3}\right|_{x=x_0} + \cdots + \frac{1}{n!}(x - x_0)^n \left.\frac{\partial^n b_k}{\partial x^n}\right|_{x=x_0} , \tag{5}$$

and

$$c_k(x) = c_k(x_0) + (x - x_0)\left.\frac{\partial c_k}{\partial x}\right|_{x=x_0} + \frac{1}{2}(x - x_0)^2 \left.\frac{\partial^2 c_k}{\partial x^2}\right|_{x=x_0} + \frac{1}{6}(x - x_0)^3 \left.\frac{\partial^3 c_k}{\partial x^3}\right|_{x=x_0} + \cdots + \frac{1}{n!}(x - x_0)^n \left.\frac{\partial^n c_k}{\partial x^n}\right|_{x=x_0} , \tag{6}$$

respectively.

10.2.3 First-Order Extended Kalman Filter

Suppose that filtered estimates $\hat{x}_{k/k}$ of x_k are desired given observations

$$z_k = c_k(x_k) + v_k , \tag{7}$$

where v_k is a measurement noise sequence. A first-order EKF for the above problem is developed below. Following the approach within [3], the nonlinear system (2) – (3) is approximated by

$$x_{k+1} = A_k x_k + B_k w_k + \mu_k , \tag{8}$$

$$y_k = C_k x_k + \pi_k , \tag{9}$$

where A_k, B_k, C_k, μ_k and π_k are found from suitable truncations of the Taylor series for each nonlinearity. From Chapter 4, a filter for the above model is given by

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k \hat{x}_{k/k-1} - \pi_k) , \tag{10}$$

$$\hat{x}_{k+1/k} = A_k \hat{x}_{k/k} + \mu_k , \tag{11}$$

"You will always define events in a manner which will validate your agreement with reality." *Steve Anthony Maraboli* where $L_k = P_{k/k-1}C_k^T\Omega_k^{-1}$ is the filter gain, in which $\Omega_k = C_kP_{k/k-1}C_k^T + R_k$, $P_{k/k} = P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1}$ and $P_{k+1/k} = A_kP_{k/k}A_k^T + B_kQ_kB_k^T$. It is common practice (see [1] – [5]) to linearise about the current conditional mean estimate, retain up to first order terms within the corresponding Taylor series and assume $B_k = b_k(\hat{x}_{k/k})$. This leads to

$$\begin{aligned} a_k(x) &\approx a_k(\hat{x}_{k/k}) + (x - \hat{x}_{k/k})^T \nabla a_k\Big|_{x=\hat{x}_{k/k}} \\ &= A_kx_k + \mu_k \end{aligned} \tag{12}$$

and

$$\begin{aligned} c_k(x_k) &\approx c_k(\hat{x}_{k/k-1}) + (x - \hat{x}_{k/k-1})^T \nabla c_k\Big|_{x=\hat{x}_{k/k-1}} \\ &= C_kx_k + \pi_k , \end{aligned} \tag{13}$$

where $A_k = \nabla a_k(x)\big|_{x=\hat{x}_{k/k}}$, $C_k = \nabla c_k(x)\big|_{x=\hat{x}_{k/k-1}}$, $\mu_k = a_k(\hat{x}_{k/k}) - A_k\hat{x}_{k/k}$ and $\pi_k = c_k(\hat{x}_{k/k-1}) - C_k\hat{x}_{k/k-1}$. Substituting for μ_k and π_k into (10) – (11) gives

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - c_k(\hat{x}_{k/k-1})) , \tag{14}$$

$$\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k}) . \tag{15}$$

Note that nonlinearities enter into the state correction (14) and prediction (15), whereas linearised matrices A_k, B_k and C_k are employed in the Riccati equation and gain calculations.

In the case of scalar states, the linearisations are $A_k = \left.\dfrac{\partial a_k}{\partial x}\right|_{x=\hat{x}_{k/k}}$ and $C_k = \left.\dfrac{\partial c_k}{\partial x}\right|_{x=\hat{x}_{k/k-1}}$. In texts on optimal filtering, the recursions (14) – (15) are either called a first-order EKF or simply an EKF, see [1] – [5]. Two higher order versions are developed below.

10.2.4 Second-Order Extended Kalman Filter

Truncating the series (1) after the second-order term and observing that $(x - \hat{x}_{k/k})^T\nabla^T$ is a scalar yields

$$\begin{aligned} a_k(x) &\approx a_k(\hat{x}_{k/k}) + (x - \hat{x}_{k/k})^T \nabla a_k\Big|_{x=\hat{x}_{k/k}} + \frac{1}{2}(x - \hat{x}_{k/k})^T \nabla^T \nabla a_k\Big|_{x=\hat{x}_{k/k}} (x - \hat{x}_{k/k}) , \\ &= a_k(\hat{x}_{k/k}) + (x - \hat{x}_{k/k})^T \nabla a_k\Big|_{x=\hat{x}_{k/k}} + \frac{1}{2}\nabla P_{k/k}\nabla^T a_k\Big|_{x=\hat{x}_{k/k}} \\ &= A_kx_k + \mu_k , \end{aligned} \tag{16}$$

"People take the longest possible paths, digress to numerous dead ends, and make all kinds of mistakes. Then historians come along and write summaries of this messy, nonlinear process and make it appear like a simple straight line." *Dean L. Kamen* where $A_k = \nabla a_k(x)\big|_{x=\hat{x}_{k/k}}$ and $\mu_k = a_k(\hat{x}_{k/k}) - A_k\hat{x}_{k/k} + \frac{1}{2}\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}}$. Similarly for the system output,

$$
\begin{aligned}
c_k(x) &\approx c_k(\hat{x}_{k/k-1}) + (x_k - \hat{x}_{k/k-1})^T \nabla c_k\big|_{x=\hat{x}_{k/k-1}} + \frac{1}{2}(x_k - \hat{x}_{k/k-1})^T \nabla^T \nabla c_k\big|_{x=\hat{x}_{k/k-1}} (x_k - \hat{x}_{k/k-1}) \\
&= c_k(\hat{x}_{k/k-1}) + (x_k - \hat{x}_{k/k-1})^T \nabla c_k\big|_{x=\hat{x}_{k/k-1}} + \frac{1}{2}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}} \\
&= C_k x_k + \pi_k ,
\end{aligned} \tag{17}
$$

where $C_k = \nabla c_k(x)\big|_{x=\hat{x}_{k/k-1}}$ and $\pi_k = c_k(\hat{x}_{k/k-1}) - C_k\hat{x}_{k/k-1} + \frac{1}{2}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}}$. Substituting for μ_k and π_k into the filtering and prediction recursions (10) – (11) yields the second-order EKF

$$
\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k\left(z_k - c_k(\hat{x}_{k/k-1}) - \frac{1}{2}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}} \right), \tag{18}
$$

$$
\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k}) + \frac{1}{2}\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}} . \tag{19}
$$

The above form is described in [2]. The further simplifications $\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}} \approx tr\left(\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}}\right)$ and $\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}} \approx tr\left(\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}}\right)$ are assumed in [4], [5].

10.2.5 Third-Order Extended Kalman Filter

Higher order EKFs can be realised just as elegantly as its predecessors. Retaining up to third-order terms within (1) results in

$$
\begin{aligned}
a_k(x) &\approx a_k(\hat{x}_{k/k}) + (x - \hat{x}_{k/k})\nabla a_k\big|_{x=\hat{x}_{k/k}} + \frac{1}{2}\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}} \\
&\quad + \frac{1}{6}\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}} (x_k - \hat{x}_{k/k})\nabla a_k\big|_{x=\hat{x}_{k/k}} \\
&= A_k x_k + \mu_k ,
\end{aligned} \tag{20}
$$

where

$$
A_k = \nabla a_k(x)\big|_{x=\hat{x}_{k/k}} + \frac{1}{6}\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}} \tag{21}
$$

"It might be a good idea if the various countries of the world would occasionally swap history books, just to see what other people are doing with the same set of facts." *William E. Vaughan* and μ_k = $a_k(\hat{x}_{k/k})$ – $A_k\hat{x}_{k/k}$ + $\frac{1}{2}\nabla P_{k/k}\nabla^T a_k\big|_{x=\hat{x}_{k/k}}$. Similarly, for the output nonlinearity it is assumed that

$$\begin{aligned} c_k(x_k) &\approx c_k(\hat{x}_{k/k-1}) + (x - \hat{x}_{k/k-1})\nabla c_k\big|_{x=\hat{x}_{k/k-1}} + \frac{1}{2}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}} \\ &+ \frac{1}{6}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}} (x_k - \hat{x}_{k/k-1})\nabla c_k\big|_{x=\hat{x}_{k/k-1}} \\ &= C_k x_k + \pi_k , \end{aligned} \tag{22}$$

where

$$C_k = \nabla c_k(x)\big|_{x=\hat{x}_{k/k-1}} + \frac{1}{6}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}} \tag{23}$$

and π_k = $c_k(\hat{x}_{k/k-1})$ – $C_k\hat{x}_{k/k-1}$ + $\frac{1}{6}\nabla P_{k/k-1}\nabla^T c_k\big|_{x=\hat{x}_{k/k-1}}$. The resulting third-order EKF is defined by (18) – (19) in which the gain is now calculated using (21) and (23).

Example 1. Consider a linear state evolution $x_{k+1} = Ax_k + w_k$, with $A = 0.5$, $w_k \in \mathbb{R}$, $Q = 0.05$, a nonlinear output mapping $y_k = \sin(x_k)$ and noisy observations $z_k = y_k + v_k$, $v_k \in \mathbb{R}$. The first-order EKF for this problem is given by

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - \sin(\hat{x}_{k/k-1})) ,$$

$$\hat{x}_{k+1/k} = A\hat{x}_{k/k} ,$$

where L_k = $P_{k/k-1}C_k^T\Omega_k^{-1}$, Ω_k = $C_kP_{k/k-1}C_k^T$ + R_k, C_k = $\cos(\hat{x}_{k/k-1})$, $P_{k/k}$ = $P_{k/k-1}$ – $P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T$ + $R_k)^{-1}C_kP_{k/k-1}$ and $P_{k+1/k}$ = $A_kP_{k/k}A_k^T$ + Q_k. The filtering step within the second-order EKF is amended to

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - \sin(\hat{x}_{k/k-1}) + \sin(\hat{x}_{k/k-1})P_{k/k-1}/2) .$$

The modified output linearisation for the third-order EKF is

$$C_k = \cos(\hat{x}_{k/k-1}) + \sin(\hat{x}_{k/k-1})P_{k/k-1}/6 .$$

Simulations were conducted in which the signal-to-noise-ratio was varied from 20 dB to 40 dB for N = 200,000 realisations of Gaussian noise sequences. The mean-square-errors exhibited by the first, second and third-order EKFs are plotted in Fig. 1. The figure demonstrates that including higher-order Taylor series terms within the filter can provide small performance improvements but the benefit diminishes with increasing measurement noise.

"No two people see the external world in exactly the same way. To every separate person a thing is what he thinks it is – in other words, not a thing, but a think." *Penelope Fitzgerald*

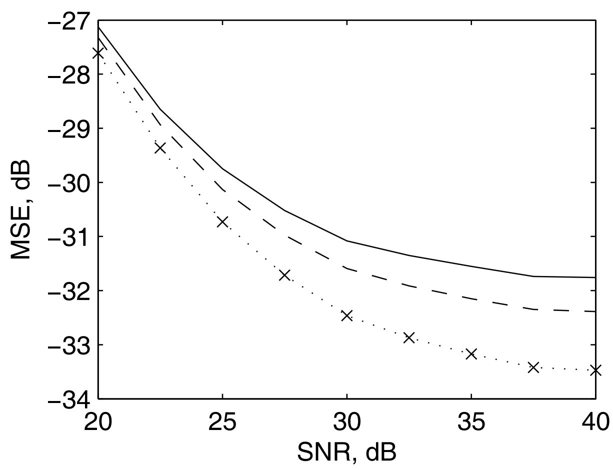


Figure 1. Mean-square-error (MSE) versus signal-to-noise-ratio (SNR) for Example 1: first-order EKF (solid line), second-order EKF (dashed line) and third-order EKF (dotted-crossed line).

10.3 The Faux Algebraic Riccati Equation Technique

10.3.1 A Nonlinear Observer

The previously-described Extended-Kalman filters arise by linearising the signal model about the current state estimate and using the linear Kalman filter to predict the next estimate. This attempts to produce a locally optimal filter, however, it is not necessarily stable because the solutions of the underlying Riccati equations are not guaranteed to be positive definite. The faux algebraic Riccati technique [6] – [10] seeks to improve on EKF performance by trading off approximate optimality for stability. The familiar structure of the EKF is retained but stability is achieved by selecting a positive definite solution to a faux Riccati equation for the gain design.

Assume that data is generated by the following signal model comprising a stable, linear state evolution together with a nonlinear output mapping

$$x_{k+1} = Ax_k + Bw_k , \tag{24}$$

$$z_k = c_k(x_k) + v_k , \tag{25}$$

where the components of $c_k(.)$ are assumed to be continuous differentiable functions. Suppose that it is desired to calculate estimates of the states from the measurements. A nonlinear observer may be constructed having the form

$$\hat{x}_{k+1/k} = A\hat{x}_k + g_k(z_k - c(\hat{x}_{k/k-1})) , \tag{26}$$

where $g_k(.)$ is a gain function to be designed. From (24) – (26), the state prediction error is given by

“The observer, when he seems to himself to be observing a stone, is really, if physics is to be believed, observing the effects of the stone upon himself.” *Bertrand Arthur William Russell*

$$\tilde{x}_{k+1/k} = A\tilde{x}_{k/k-1} - g_k(\varepsilon_k) + w_k \, , \tag{27}$$

where $\tilde{x}_k = x_k - \hat{x}_{k/k-1}$ and $\varepsilon_k = z_k - c(\hat{x}_{k/k-1})$. The Taylor series expansion of $c_k(.)$ to first order terms leads to $\varepsilon_k \approx C_k\tilde{x}_{k/k-1} + v_k$, where $C_k = \nabla c_k(x)\big|_{x=\hat{x}_{k/k-1}}$. The objective here is to design $g_k(\varepsilon_k)$ to be a linear function of $\tilde{x}_{k/k-1}$ to first order terms. It will be shown that for certain classes of problems, this objective can be achieved by a suitable choice of a linear bounded matrix function of the states D_k, resulting in the time-varying gain function $g_k(\varepsilon_k) = K_kD_k\varepsilon_k$, where K_k is a gain matrix of appropriate dimension. For example, consider $x_k \in \mathbb{R}^n$ and $z_k \in \mathbb{R}^m$, which yield $\varepsilon_k \in \mathbb{R}^m$ and $C_k \in \mathbb{R}^{m\times n}$. Suppose that a linearisation $D_k \in \mathbb{R}^{p\times m}$ can be found so that $\bar{C}_k = D_kC_k \in \mathbb{R}^{p\times m}$ possesses approximately constant terms. Then the locally linearised error (27) may be written as

$$\tilde{x}_{k+1/k} = (A - K_k\bar{C}_k)\tilde{x}_{k/k-1} - K_kD_kv_k + w_k \, . \tag{28}$$

If $|\lambda_i(A)| < 1$, $i = 1 \ldots n$, and if the pair $(A, \bar{C}_k)$ is completely observable, then the asymptotic stability of (28) can be guaranteed by selecting the gain such that $|\lambda_i(A - K_k\bar{C}_k)| < 1$. A method for selecting the gain is described below.

10.3.2 Gain Selection

From (28), an approximate equation for the error covariance $P_{k/k-1} = E\{\tilde{x}_{k/k-1}\tilde{x}^T_{k/k-1}\}$ is

$$P_{k+1/k} = (A - K_k\bar{C}_k)P_{k/k-1}(A - K_k\bar{C}_k)^T + K_kD_kRD_k^TK_k^T + Q \, , \tag{29}$$

which can be written as

$$P_{k/k} = P_{k/k-1} - P_{k/k-1}\bar{C}_k^T(\bar{C}_kP_{k/k-1}\bar{C}_k^T + D_kRD_k^T)^{-1}\bar{C}_kP_{k/k-1} \, , \tag{30}$$

$$P_{k+1/k} = AP_{k/k}A + Q \, . \tag{31}$$

In an EKF for the above problem, the gain is obtained by solving the above Riccati difference equation and calculating

$$K_k = P_{k/k-1}\bar{C}_k^T(\bar{C}_kP_{k/k-1}\bar{C}_k^T + D_kRD_k^T)^{-1} \, . \tag{32}$$

The faux algebraic Riccati equation approach [6] – [10] is motivated by connections between Riccati difference equation and algebraic Riccati equation solutions. Indeed, it is noted for some nonlinear problems that the gains can converge to a steady-state matrix [3]. This technique is also known as 'covariance setting'. Following the approach of [10], the Riccati difference equation (30) may be masqueraded by the faux algebraic Riccati equation

$$\Sigma_k = \Sigma_k - \Sigma_k\bar{C}_k^T(\bar{C}_k\Sigma_k\bar{C}_k^T + D_kRD_k^T)^{-1}\bar{C}_k\Sigma_k \, . \tag{33}$$

"The universe as we know it is a joint product of the observer and the observed." *Pierre Teilhard De Chardin*

That is, rather than solve (30), an arbitrary positive definite solution Σ_k is assumed instead and then the gain at each time k is calculated from (31) – (32) using Σ_k in place of $P_{k/k-1}$.

10.3.3 Tracking Multiple Signals

Consider the problem of tracking two frequency or phase modulated signals which may be modelled by equation (34), where $a_k^{(1)}$, $a_k^{(2)}$, $\omega_k^{(1)}$, $\omega_k^{(2)}$, $\phi_k^{(1)}$, $\phi_k^{(2)}$, $\mu_a^{(1)}$, $\mu_a^{(2)}$, $\mu_\omega^{(1)}$, $\mu_\omega^{(2)}$ $\in \mathbb{R}$ and $w_k^{(1)}$, … $w_k^{(6)}$ $\in \mathbb{R}$ are zero-mean, uncorrelated, white processes with covariance Q = diag($\sigma^2_{w^{(1)}}, \ldots, \sigma^2_{w^{(6)}}$). The states $a_k^{(i)}$, $\omega_k^{(i)}$ and $\phi_k^{(i)}$, i = 1, 2, represent the signals' instantaneous amplitude, frequency and phase components, respectively.

$$\begin{bmatrix} a_{k+1}^{(1)} \\ \omega_{k+1}^{(1)} \\ \phi_{k+1}^{(1)} \\ a_{k+1}^{(2)} \\ \omega_{k+1}^{(2)} \\ \phi_{k+1}^{(2)} \end{bmatrix} = \begin{bmatrix} \mu_a^{(1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_\omega^{(1)} & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_a^{(2)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu_\omega^{(2)} & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_k^{(1)} \\ \omega_k^{(1)} \\ \phi_k^{(1)} \\ a_k^{(2)} \\ \omega_k^{(2)} \\ \phi_k^{(2)} \end{bmatrix} + \begin{bmatrix} w_k^{(1)} \\ w_k^{(2)} \\ w_k^{(3)} \\ w_k^{(4)} \\ w_k^{(5)} \\ w_k^{(6)} \end{bmatrix}. \quad (34)$$

Let

$$\begin{bmatrix} z_k^{(1)} \\ z_k^{(2)} \\ z_k^{(3)} \\ z_k^{(4)} \end{bmatrix} = \begin{bmatrix} a_k^{(1)} \cos \phi_k^{(1)} \\ a_k^{(1)} \sin \phi_k^{(1)} \\ a_k^{(2)} \cos \phi_k^{(2)} \\ a_k^{(2)} \sin \phi_k^{(2)} \end{bmatrix} + \begin{bmatrix} v_k^{(1)} \\ v_k^{(2)} \\ v_k^{(3)} \\ v_k^{(4)} \end{bmatrix} \quad (35)$$

denote the complex baseband observations, where $v_k^{(1)}$, …, $v_k^{(4)}$ $\in \mathbb{R}$ are zero-mean, uncorrelated, white processes with covariance R = $diag(\sigma^2_{v^{(1)}}, \ldots, \sigma^2_{v^{(4)}})$. Expanding the prediction error to linear terms yields $C_k = [C_k^{(1)} \;\; C_k^{(2)}]$, where

$$C_k^{(i)} = \begin{bmatrix} \cos \hat{\phi}_{k/k-1}^{(i)} & 0 & -\hat{a}_{k/k-1}^{(i)} \sin \hat{\phi}_{k/k-1}^{(i)} \\ \sin \hat{\phi}_{k/k-1}^{(i)} & 0 & \hat{a}_{k/k-1}^{(i)} \cos \hat{\phi}_{k/k-1}^{(i)} \end{bmatrix}.$$

This form suggests the choice $D_k = \begin{bmatrix} D_k^{(1)} \\ D_k^{(2)} \end{bmatrix}$, where

$$D_k^{(i)} = \begin{bmatrix} \cos \hat{\phi}_{k/k-1}^{(i)} & \sin \hat{\phi}_{k/k-1}^{(i)} \\ -\sin \hat{\phi}_{k/k-1}^{(i)} / \hat{a}_{k/k-1}^{(i)} & \cos \hat{\phi}_{k/k-1}^{(i)} / \hat{a}_{k/k-1}^{(i)} \end{bmatrix}.$$

"If you haven't found something strange during the day, it hasn't been much of a day." *John Archibald Wheeler*

In the multiple signal case, the linearization $\bar{C}_k = D_k C_k$ does not result in perfect decoupling. While the diagonal blocks reduce to $\bar{C}_k^{(i,i)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, the off-diagonal blocks are

$$\bar{C}_k^{(i,j)} = \begin{bmatrix} \cos(\hat{\phi}_{k/k-1}^{(i)} - \hat{\phi}_{k/k-1}^{(j)}) & 0 & \hat{a}_{k/k-1}^{(j)} \sin(\hat{\phi}_{k/k-1}^{(i)} - \hat{\phi}_{k/k-1}^{(j)}) \\ -\frac{1}{\hat{a}_{k/k-1}^{(i)}} \cos(\hat{\phi}_{k/k-1}^{(i)} - \hat{\phi}_{k/k-1}^{(j)}) & 0 & \frac{\hat{a}_{k/k-1}^{(j)}}{\hat{a}_{k/k-1}^{(i)}} \cos(\hat{\phi}_{k/k-1}^{(i)} - \hat{\phi}_{k/k-1}^{(j)}) \end{bmatrix}.$$

Assuming a symmetric positive definite solution to (33) of the form $\Sigma_k = \begin{bmatrix} \Sigma_k^a & 0 & 0 \\ 0 & \Sigma_k^{\omega} & \Sigma_k^{\omega\phi} \\ 0 & \Sigma_k^{\phi\omega} & \Sigma_k^{\phi} \end{bmatrix}$, with Σ_k^a, Σ_k^{ω}, $\Sigma_k^{\omega\phi}$, $\Sigma_k^{\phi} \in \mathbb{R}$ and choosing the gains according to (32) yields $K_k = \begin{bmatrix} K_k^a & 0 \\ 0 & K_k^{\omega} \\ 0 & K_k^{\phi} \end{bmatrix}$, where $K_k^a = \Sigma_k^a(\Sigma_k^a + \sigma_v^2)^{-1}$, $K_k^{\omega} = \Sigma_k^{\omega\phi}(\Sigma_k^{\phi} + \sigma_v^2 \hat{a}_{k/k}^{-2})^{-1}$ and $K_k^{\phi} = \Sigma_k^{\phi}(\Sigma_k^{\phi} + \sigma_v^2 \hat{a}_{k/k}^{-2})^{-1}$. The nonlinear observer then becomes

$$\hat{a}_{k/k}^{(i)} = \hat{a}_{k/k-1}^{(i)} + \Sigma_k^a (z_k^{(1)} \cos \hat{\phi}_{k/k-1}^{(i)} + z_k^{(2)} \sin \hat{\phi}_{k/k-1}^{(i)})(\Sigma_k^a + \sigma_v^2)^{-1},$$

$$\hat{\omega}_{k/k}^{(i)} = \hat{\omega}_{k/k-1}^{(i)} + \Sigma_k^{\omega\phi} (z_k^{(1)} \cos \hat{\phi}_{k/k-1}^{(i)} + z_k^{(2)} \sin \hat{\phi}_{k/k-1}^{(i)})(\hat{a}_{k/k-1}^{(i)} \Sigma_k^{\phi} + \sigma_v^2 / a_{k/k-1}^{(i)})^{-1},$$

$$\hat{\phi}_{k/k}^{(i)} = \hat{\phi}_{k/k-1}^{(i)} + \Sigma_k^{\phi} (z_k^{(1)} \cos \hat{\phi}_{k/k-1}^{(i)} + z_k^{(2)} \sin \hat{\phi}_{k/k-1}^{(i)})(\hat{a}_{k/k-1}^{(i)} \Sigma_k^{\phi} + \sigma_v^2 / a_{k/k-1}^{(i)})^{-1}.$$

10.3.4 Stability Conditions

In order to establish conditions for the error system (28) to be asymptotically stable, the problem is recast in a passivity framework as follows. Let $w = \begin{bmatrix} w_k^{(1)} \\ w_k^{(2)} \\ \vdots \\ w_k^{(m)} \end{bmatrix}$, $e = \begin{bmatrix} e_k^{(1)} \\ e_k^{(2)} \\ \vdots \\ e_k^{(m)} \end{bmatrix} \in \mathbb{R}^m$.

Consider the configuration of Fig. 2, in which there is a cascade of a stable linear system $\mathcal{G}$ and a nonlinear function matrix $\gamma(.)$ acting on e. It follows from the figure that

$$e = w - \mathcal{G}\gamma(e). \quad (36)$$

Let Δ_f denote a forward difference operator with $\Delta_f e_k = e_k^{(i)} - e_{k-1}^{(i)}$. It is assumed that $\gamma(.)$ satisfies some sector conditions which may be interpreted as bounds existing on the slope of the components of $\gamma(.)$; see Theorem 14, p. 7 of [11].

"Discovery consists of seeing what everyone has seen and thinking what nobody has thought." *Albert Szent-Görgyi*

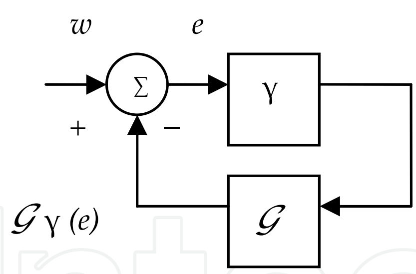


Figure 2. Nonlinear error system configuration.

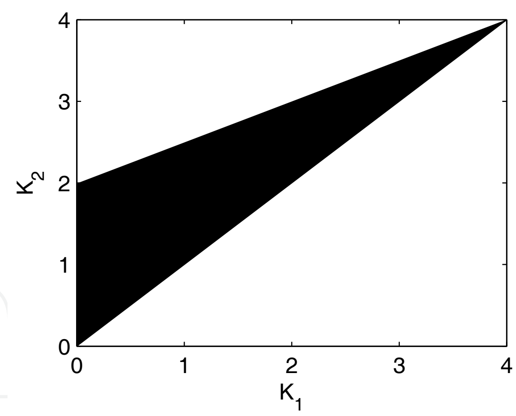


Figure 3. Stable gain space for Example 2.

Lemma 1 [10]: *Consider the system (36), where w, e* $\in \mathbb{R}^m$ *. Suppose that* $\gamma(.)$ *consists of m identical, noninteracting nonlinearities, with* $\gamma(e_k^{(i)})$ *monotonically increasing in the sector [0,β], β ≥ 0, β ∈* $\mathbb{R}$ *, that is,*

$$0 \le \gamma(e_k^{(i)}) / e_k^{(i)} \le \beta \tag{37}$$

for all $e_k^{(i)} \in \mathbb{R}$ *,* $e_k^{(i)} \neq 0$*. Assume that* $\mathcal{G}$ *is a causal, stable, finite-gain, time-invariant map* $\mathbb{R}^m \to \mathbb{R}^m$ *, having a z-transform G(z), which is bounded on the unit circle. Let I denote an* $m \times m$ *identity matrix. Suppose that for some q > 0, q ∈* $\mathbb{R}$ *, there exists a* $\delta \in \mathbb{R}$ *, such that*

$$\left\langle (G(z) + q\Delta_f G(z) + I\beta^{-1})e, e \right\rangle \ge \delta \left\langle e, e \right\rangle \tag{38}$$

for all $e_k^{(i)} \in \mathbb{R}$ *. Under these conditions* $w \in \ell_2$ *implies e,* $\gamma(e_k^{(i)}) \in \ell_2$ *.*

Proof: *From (36),* $\Delta_f w = \Delta_f e + \Delta_f G(z)\gamma(e)$ *and* $w + q\Delta_f w = (G(z) + q\Delta_f G(z) + I\beta^{-1})\gamma(e) + e - I\beta^{-1}\gamma(e) + e - I\beta^{-1}\gamma(e) + q\Delta_f e$ *. Then*

$$\begin{aligned} \left\langle w + q\Delta_f w, \gamma(e) \right\rangle &\ge \left\langle e - I\beta^{-1}\gamma(e), \gamma(e) \right\rangle + \left\langle q\Delta_f e), \gamma(e) \right\rangle \\ &\quad + \left\langle (G(z) + q\Delta_f G(z) + I\beta^{-1})\gamma(e), \gamma(e) \right\rangle . \end{aligned} \tag{39}$$

Consider the first term on the right hand side of (39). Since the $\gamma(e)$ *consists of noninteracting nonlinearities,* $\left\langle \gamma(e), e \right\rangle = \sum_{i=1}^{m} \left\langle \gamma(e^{(i)}), e^{(i)} \right\rangle$ *and* $\left\langle e - I\beta^{-1}\gamma(e), \gamma(e) \right\rangle = \sum_{i=1}^{m} \left\langle e^{(i)} - \gamma(e^{(i)}) I\beta^{-1}, e^{(i)} \right\rangle \ge 0$*. Using the approach of [11] together with the sector conditions on the identical noninteracting nonlinearities (37), it can be shown that expanding out the second term of (39) yields* $\left\langle \Delta_f e, \gamma(e) \right\rangle \ge$

"The intelligent man finds almost everything ridiculous, the sensible man hardly anything." *Johann Wolfgang von Goethe*

0. Using $\left\|\Delta_f w\right\|_2 \le 2\left\|w\right\|_2$ *(from p. 192 of [11]), the Schwartz inequality and the triangle inequality, it can be shown that*

$$\left\langle w + q\nabla w, \gamma(e)\right\rangle \le (1+2q)\left\|w\right\|_2 . \tag{40}$$

It follows from (38) – (40) that $\left\|\gamma(e)\right\|_2^2 \le (1 + 2q)\delta^{-1}\left\|w\right\|_2$*; hence* $\gamma(e_k^{(i)}) \in \ell_2$*. Since the gain of G(z) is finite, it also follows that* $G(z)\gamma(e_k^{(i)}) \in \ell_2$*.* □

If $G(z)$ is stable and bounded on the unit circle, then the test condition (38) becomes

$$\lambda_{\min}\{I + q(I - z^{-1}I)(G(z) + G^H(z)) + \beta^{-1}\} \ge \delta , \tag{41}$$

see pp. 175 and 194 of [11].

10.3.5 Applications

Example 2 [10]. Consider a unity-amplitude frequency modulated (FM) signal modelled as $\omega_{k+1} = \mu_\omega \omega_k + w_k$, $\phi_{k+1} = \phi_k + \omega_k$, $z_k^{(1)} = \cos(\phi_k) + v_k^{(1)}$ and $z_k^{(2)} = \sin(\phi_k) + v_k^{(2)}$. The error system for an FM demodulator may be written as

$$\begin{bmatrix} \tilde{\omega}_{k+1} \\ \tilde{\phi}_{k+1} \end{bmatrix} = \begin{bmatrix} \mu_\omega & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\omega}_k \\ \tilde{\phi}_k \end{bmatrix} - \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \sin(\tilde{\phi}_k) + w_k \tag{42}$$

for gains K_1, $K_2 \in \mathbb{R}$ to be designed. In view of the form (36), the above error system is reformatted as

$$\begin{bmatrix} \tilde{\omega}_{k+1} \\ \tilde{\phi}_{k+1} \end{bmatrix} = \begin{bmatrix} \mu_\omega & -K_1 \\ 1 & 1-K_2 \end{bmatrix} \begin{bmatrix} \tilde{\omega}_k \\ \tilde{\phi}_k \end{bmatrix} + \gamma\left(\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\omega}_k \\ \tilde{\phi}_k \end{bmatrix} \right) + w_k , \tag{43}$$

where $\gamma(x) = x - \sin(x)$. The z-transform of the linear part of (43) is $G(z) = (K_2 z + K_2 + \mu_\omega K_1)(z^2 + (K_2 - 1 - \mu_\omega)z + K_1 + 1 - \mu_\omega K_2)^{-1}$. The nonlinearity satisfies the sector condition (37) for $\beta = 1.22$. Candidate gains may be assessed by checking that $G(z)$ is stable and the test condition (41). The stable gain space calculated for the case of $\mu_\omega = 0.9$ is plotted in Fig. 3. The gains are required to lie within the shaded region of the plot for the error system (42) to be asymptotically stable.

"He that does not offend cannot be honest." *Thomas Paine*

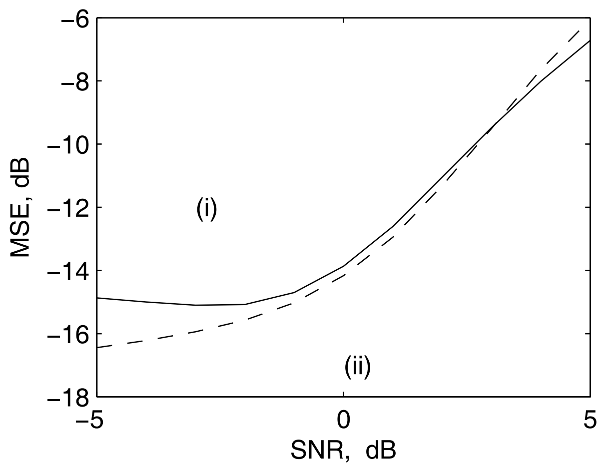


Figure 4. Demodulation performance for Example 2: (i) EKF and (ii) Nonlinear observer.

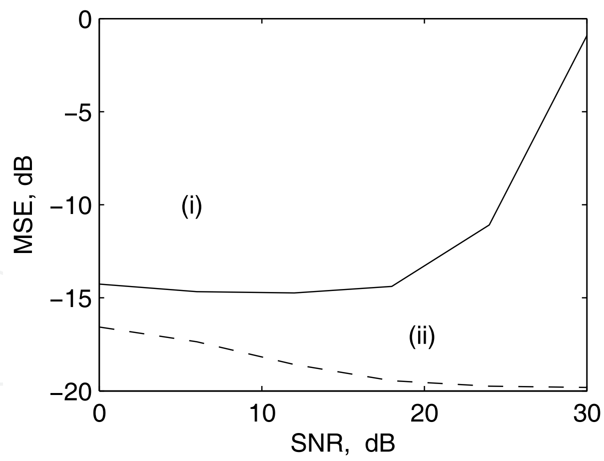


Figure 5. Demodulation performance for Example 3: (i) EKF and (ii) Nonlinear observer.

A speech utterance, namely, the phrase "Matlab is number one", was sampled at 8 kHz and used to synthesize a unity-amplitude FM signal. An EKF demodulator was constructed for the above model with σ_w^2 = 0.02. In a nonlinear observer design it was found that suitable parameter choices were Σ_k = $\begin{bmatrix} 0.001 & 0.08 \\ 0.08 & 0.7 \end{bmatrix}$. The nonlinear observer gains were censored at each time k according to the stable gain space of Fig. 3. The results of a simulation study using 100 realisations of Gaussian measurement noise sequences are shown in Fig. 4. The figure demonstrates that enforcing stability can be beneficial at low SNR, at the cost of degraded high-SNR performance.

Example 3 [10]. Suppose that there are two superimposed FM signals present in the same frequency channel. Neglecting observation noise, a suitable approximation of the demodulator error system in the form (36) is given by

$$\begin{bmatrix} \tilde{\omega}_{k+1}^{(1)} \\ \tilde{\phi}_{k+1}^{(1)} \\ \tilde{\omega}_{k+1}^{(2)} \\ \tilde{\phi}_{k+1}^{(2)} \end{bmatrix} = (A - K_k\bar{C}) \begin{bmatrix} \tilde{\omega}_{k}^{(1)} \\ \tilde{\phi}_{k}^{(1)} \\ \tilde{\omega}_{k}^{(2)} \\ \tilde{\phi}_{k}^{(2)} \end{bmatrix} - K_k \begin{bmatrix} \sin(\tilde{\phi}_k^{(1)}) - \tilde{\phi}_k^{(1)} \\ \sin(\tilde{\phi}_k^{(2)}) - \tilde{\phi}_k^{(2)} \end{bmatrix}, \tag{44}$$

where A = diag($A^{(1)}$, $A^{(1)}$), $A^{(1)}$ = $\begin{bmatrix} \mu_\omega & 0 \\ 1 & 1 \end{bmatrix}$, $\bar{C}$ = $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. The linear part of (44) may be written as $G(z) = \bar{C}(zI - (A - K_k\bar{C}))^{-1}K_k$. Two 8-kHz speech utterances, "Matlab is number one" and "Number one is Matlab", centred at ±0.25 rad/s, were used to synthesize two superimposed unity-amplitude FM signals. Simulations were conducted using 100 realisations of Gaussian measurement noise sequences. The test condition (41) was

"To avoid criticism, do nothing, say nothing, be nothing." *Elbert Hubbard* evaluated at each time k for the above parameter values with β = 1.2, q = 0.001, δ = 0.82 and used to censor the gains. The resulting co-channel demodulation performance is shown in Fig. 5. It can be seen that the nonlinear observer significantly outperforms the EKF at high SNR.

Two mechanisms have been observed for occurrence of outliers or faults within the co-channel demodulators. Firstly errors can occur in the state attribution, that is, there is correct tracking of some component speech message segments but the tracks are inconsistently associated with the individual signals. This is illustrated by the example frequency estimate tracks shown in Figs. 6 and 7. The solid and dashed lines in the figures indicate two sample co-channel frequency tracks. Secondly, the phase unwrapping can be erroneous so that the frequency tracks bear no resemblance to the underlying messages. These faults can occur without any significant deterioration in the error residual.

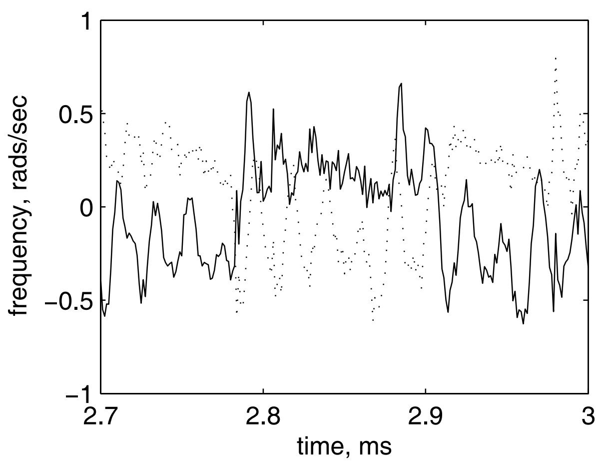


Figure 6. Sample EKF frequency tracks for Example 3.

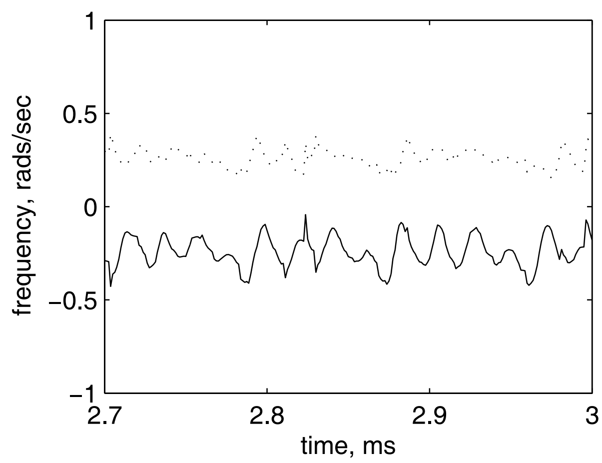


Figure 7. Sample Nonlinear observer frequency tracks for Example 3.

The EKF demodulator is observed to be increasingly fault prone at higher SNR. This arises because lower SNR designs possess narrower bandwidths and so are less sensitive to nearby frequency components. The figures also illustrate the trade-off between stability and optimality. In particular, it can be seen from Fig. 6, that the sample EKF speech estimates exhibit faults in the state attribution. This contrasts with Fig. 7, where the nonlinear observer's estimates exhibit stable state attribution at the cost of degraded speech fidelity.

10.4 Robust Extended Kalman Filtering

10.4.1 Nonlinear Problem Statement

Consider again the nonlinear, discrete-time signal model (2), (7). It is shown below that the H_∞ techniques of Chapter 9 can be used to recast nonlinear filtering problems into a model uncertainty setting. The following discussion attends to state estimation, that is, $C_{1,k} = I$ is assumed within the problem and solution presented in Section 9.3.2.

"You have enemies? Good. That means you've stood up for something, sometime in your life." *Winston Churchill*

The Taylor series expansions of the nonlinear functions $a_k(.)$, $b_k(.)$ and $c_k(.)$ about filtered and predicted estimates $\hat{x}_{k/k}$ and $\hat{x}_{k/k-1}$ may be written as

$$a_k(x_k) = a_k(\hat{x}_{k/k}) + \nabla a_k(\hat{x}_{k/k})(x_k - \hat{x}_{k/k}) + \Delta_1(\tilde{x}_{k/k}) , \tag{45}$$

$$b_k(x_k) = b_k(\hat{x}_{k/k}) + \Delta_2(\tilde{x}_{k/k}) , \tag{46}$$

$$c_k(x_k) = c_k(\hat{x}_{k/k-1}) + \nabla c_k(\hat{x}_{k/k-1})(x_k - \hat{x}_{k/k-1}) + \Delta_3(\tilde{x}_{k/k-1}) , \tag{47}$$

where $\Delta_1(.)$, $\Delta_2(.)$, $\Delta_3(.)$ are uncertainties that account for the higher order terms, $\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$ and $\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$. It is assumed that $\Delta_1(.)$, $\Delta_2(.)$ and $\Delta_3(.)$ are continuous operators mapping $\ell_2 \to \ell_2$, with H_∞ norms bounded by δ_1, δ_2 and δ_3, respectively.

Substituting (45) – (47) into the nonlinear system (2), (7) gives the linearised system

$$x_{k+1} = A_k x_k + B_k w_k + \mu_k + \Delta_1(\tilde{x}_{k/k}) + \Delta_2(\tilde{x}_{k/k}) w_k , \tag{48}$$

$$z_k = C_k x_k + \pi_k + \Delta_3(\tilde{x}_{k/k-1}) + v_k , \tag{49}$$

where $A_k = \nabla a_k(x)\big|_{x=\hat{x}_{k/k}}$, $C_k = \nabla c_k(x)\big|_{x=\hat{x}_{k/k-1}}$, $\mu_k = a_k(\hat{x}_{k/k}) - A_k\hat{x}_{k/k}$ and $\pi_k = c_k(\hat{x}_{k/k-1}) - C_k\hat{x}_{k/k-1}$.

Note that the first-order EKF for the above system arises by setting the uncertainties $\Delta_1(.)$, $\Delta_2(.)$ and $\Delta_3(.)$ to zero as

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - c_k(\hat{x}_{k/k-1})) , \tag{50}$$

$$\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k}) , \tag{51}$$

$$L_k = P_{k/k-1}C_k^T(C_k P_{k/k-1} C_k^T + R_k)^{-1} , \tag{52}$$

$$P_{k/k} = P_{k/k-1} - P_{k/k-1}C_k^T(C_k P_{k/k-1} C_k^T + R_k)^{-1} P_{k/k-1} C_k , \tag{53}$$

$$P_{k+1/k} = A_k P_{k/k-1} A_k^T + B_k Q_k B_k^T . \tag{54}$$

10.4.2 Robust Solution

Following the approach in Chapter 9, instead of addressing the problem (48) – (49) which possesses uncertainties, an auxiliary H_∞ problem is defined as

$$x_{k+1} = A_k x_k + B_k w_k + \mu_k + s_k , \tag{55}$$

$$z_k = C_k x_k + \pi_k + v_k + t_k , \tag{56}$$

"Fight the good fight." *Timothy 4:7*

$$\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}, \tag{57}$$

where $s_k = \Delta_1(\tilde{x}_{k/k}) + \Delta_2(\tilde{x}_{k/k})w_k$ and $t_k = \Delta_3\tilde{x}_{k/k} \approx \Delta_3\tilde{x}_{k/k-1}$ are additional exogenous inputs satisfying

$$\|s_k\|_2^2 \le \delta_1^2 \|\tilde{x}_{k/k}\|_2^2 + \delta_2^2 \|w_k\|_2^2, \tag{58}$$

$$\|t_k\|_2^2 \le \delta_3^2 \|\tilde{x}_{k/k}\|_2^2 \le \delta_3^2 \|\tilde{x}_{k/k-1}\|_2^2. \tag{59}$$

A sufficient solution to the auxiliary H∞ problem (55) – (57) can be obtained by solving another problem in which w_k and v_k are scaled in lieu of the additional inputs s_k and r_k. The scaled H∞ problem is defined by

$$x_{k+1} = A_k x_k + B_k c_w w_k + \mu_k, \tag{60}$$

$$z_k = C_k x_k + c_v v_k + \pi_k, \tag{61}$$

$$\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}, \tag{62}$$

where c_w, $c_v \in \mathbb{R}$ are to be found.

Lemma 2 [12]: *The solution of the H∞ problem (60) – (62), where v_k is scaled by*

$$c_v^2 = 1 - \gamma^2\delta_1^2 - \gamma^2\delta_3^2, \tag{63}$$

and w_k is scaled by

$$c_w^2 = c_v^2(1 + \delta_2^2)^{-1}, \tag{64}$$

is sufficient for the solution of the auxiliary H∞ problem (55) – (57).

Proof: *If the H∞ problem (50) – (52) has been solved then there exists a $\gamma \ne 0$ such that*

$$\|\tilde{x}_{k/k}\|_2^2 \le \gamma^2(\|w_k\|_2^2 + \|s_k\|_2^2 + \|t_k\|_2^2 + \|v_k\|_2^2)$$

$$\le \gamma^2(\|w_k\|_2^2 + \delta_1^2\|\tilde{x}_{k/k}\|_2^2 + \delta_2^2\|w_k\|_2^2 + \delta_3^2\|\tilde{x}_{k/k}\|_2^2 + \|v_k\|_2^2),$$

which implies

$$(1 - \gamma^2\delta_1^2 - \gamma^2\delta_3^2)\|\tilde{x}_{k/k}\|_2^2 \le \gamma^2((1 + \delta_2^2)\|w_k\|_2^2 + \|v_k\|_2^2)$$

"You can't wait for inspiration. You have to go after it with a club." *Jack London* *and*

$$\|\tilde{x}_{k/k}\|_2^2 \le \gamma^2 (c_w^{-2} \|w_k\|_2^2 + c_v^{-2} \|v_k\|_2^2) .$$ □

The robust first-order extended Kalman filter for state estimation is given by (50) – (52),

$$P_{k/k} = P_{k/k-1} - P_{k/k-1} \begin{bmatrix} I & C_k^T \end{bmatrix} \begin{bmatrix} P_{k/k-1} - \gamma^2 I & P_{k/k-1} C_k^T \\ C_k P_{k/k-1} & R_k + C_k P_{k/k-1} C_k^T \end{bmatrix}^{-1} \begin{bmatrix} I \\ C_k \end{bmatrix} P_{k/k-1}$$

and (54). As discussed in Chapter 9, a search is required for a minimum γ such that $\begin{bmatrix} P_{k/k-1} - \gamma^2 I & P_{k/k-1} C_k^T \\ C_k P_{k/k-1} & R_k + C_k P_{k/k-1} C_k^T \end{bmatrix} > 0$ and $P_{k/k-1} > 0$ over $k \in [1, N]$. An illustration is provided below.

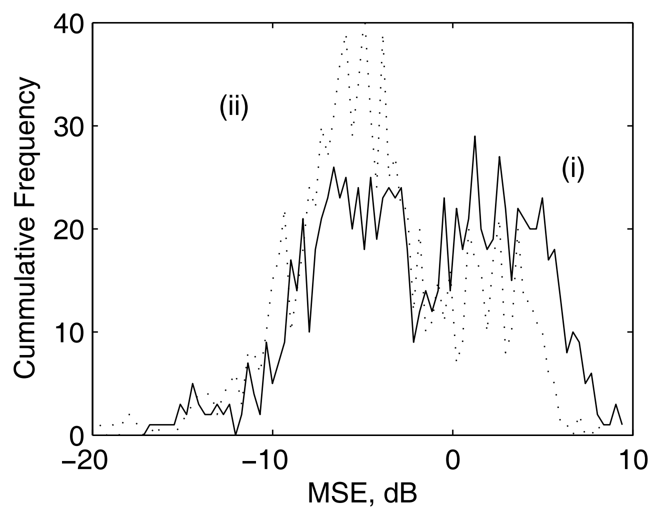


Figure 8. Histogram of demodulator mean-square-error for Example 4: (i) first-order EKF (solid line) and first-order robust EKF (dotted line).

Example 4 [12]. Suppose that an FM signal is generated by[17]

$$\omega_{k+1} = \mu_\omega \omega_k + w_k , \tag{65}$$

$$\phi_{k+1} = \arctan(\mu_\phi \phi_k + \omega_k) , \tag{66}$$

$$z_k^{(1)} = \cos(\phi_k) + v_k^{(1)} , \tag{67}$$

$$z_k^{(2)} = \sin(\phi_k) + v_k^{(2)} . \tag{68}$$

The objective is to construct an FM demodulator that produces estimates of the frequency message ω_k from the noisy in-phase and quadrature measurements $z_k^{(1)}$ and $z_k^{(2)}$, respectively. Simulations were conducted with $\mu_\omega = 0.9$, $\mu_\phi = 0.99$ and $\sigma^2_{v^{(1)}} = \sigma^2_{v^{(2)}} = 0.001$. It was found for $\sigma_w^2 < 0.1$, where the state behaviour is almost linear, a robust EKF does not

[17] "Happy is he who gets to know the reasons for things." *Virgil*

improve on the EKF. However, when σ_w^2 = 1, the problem is substantially nonlinear and a performance benefit can be observed. A robust EKF demodulator was designed with

$$x_k = \begin{bmatrix} \phi_k \\ \omega_k \end{bmatrix},\ A_k = \begin{bmatrix} \dfrac{\mu_\phi}{(\mu_\phi \hat{\phi}_{k/k} + \hat{\omega}_{k/k})^2 + 1} & \dfrac{\mu_\phi}{(\mu_\phi \hat{\phi}_{k/k} + \hat{\omega}_{k/k})^2 + 1} \\ 0 & \mu_\omega \end{bmatrix},\ C_k = \begin{bmatrix} -\sin(\hat{\phi}_{k/k-1}) & 0 \\ \cos(\hat{\phi}_{k/k-1}) & 0 \end{bmatrix},$$

δ_1 = 0.1, δ_2 = 4.5 and δ_3 = 0.001. It was found that γ = 1.38 was sufficient for $P_{k/k-1}$ of the above Riccati difference equation to always be positive definite. A histogram of the observed frequency estimation error is shown in Fig. 8, which demonstrates that the robust demodulator provides improved mean-square-error performance. For sufficiently large σ_w^2, the output of the above model will resemble a digital signal, in which case a detector may outperform a demodulator.

10.5 Nonlinear Smoothing

10.5.1 Approximate Minimum-Variance Smoother

Consider again a nonlinear estimation problem where $x_{k+1} = a_k(x_k) + B_k w_k$, $z_k = c_k(x_k) + v_k$, with $x_k \in \mathbb{R}$, in which the nonlinearities $a_k(.)$, $c_k(.)$ are assumed to be smooth, differentiable functions of appropriate dimension. The linearisations akin to Extended Kalman filtering may be applied within the smoothers described in Chapter 7 in the pursuit of performance improvement. The fixed-lag, Fraser-Potter and Rauch-Tung-Striebel smoother recursions are easier to apply as they are less complex. The application of the minimum-variance smoother can yield approximately optimal estimates when the problem becomes linear, provided that the underlying assumptions are correct.

Procedure 1. An approximate minimum-variance smoother for output estimation can be implemented via the following three-step procedure.

Step 1. Operate

$$\alpha_k = -\Omega_k^{1/2}(z_k - c_k(\hat{x}_{k/k-1})), \tag{69}$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - c_k(\hat{x}_{k/k-1})), \tag{70}$$

$$\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k}), \tag{71}$$

on the measurement z_k, where $L_k = P_{k/k-1} C_k^T \Omega_k^{-1}$,

$$\Omega_k = C_k P_{k/k-1} C_k^T + R_k,$$

$$P_{k/k} = P_{k/k-1} - P_{k/k-1} C_k^T \Omega_k^{-1} C_k P_{k/k-1},$$

$$P_{k+1/k} = A_k P_{k/k} A_k^T + B_k Q_k B_k^T, \tag{72}$$

$$A_k = \left.\frac{\partial a_k}{\partial x}\right|_{x=\hat{x}_{k/k}} \text{ and } C_k = \left.\frac{\partial c_k}{\partial x}\right|_{x=\hat{x}_{k/k-1}}.$$

"You can recognize a pioneer by the arrows in his back" *Beverly Rubik*

Step 2. Operate (69) – (71) on the time-reversed transpose of α_k. Then take the time-reversed transpose of the result to obtain β_k.

Step 3. Calculate the smoothed output estimate from

$$\hat{y}_{k/N} = z_k - R_k \beta_k . \quad (73)$$

10.5.2 Robust Smoother

From the arguments within Chapter 9, a smoother that is robust to uncertain w_k and v_k can be realised by replacing the error covariance correction (72) by

$$P_{k/k} = P_{k/k-1} - P_{k/k-1}\begin{bmatrix} C_k^T & C_k^T \end{bmatrix}\begin{bmatrix} C_k P_{k/k-1} C_k^T - \gamma^2 I & C_k P_{k/k-1} C_k^T \\ C_k P_{k/k-1} C_k^T & R_k + C_k P_{k/k-1} C_k^T \end{bmatrix}^{-1}\begin{bmatrix} C_k \\ C_k \end{bmatrix} P_{k/k-1}$$

within Procedure 1. As discussed in Chapter 9, a search for a minimum γ such that $\begin{bmatrix} C_k P_{k/k-1} C_k^T - \gamma^2 I & C_k P_{k/k-1} C_k^T \\ C_k P_{k/k-1} C_k^T & R_k + C_k P_{k/k-1} C_k^T \end{bmatrix} > 0$ and $P_{k/k-1} > 0$ over $k \in [1, N]$ is desired.

10.5.3 Application

Returning to the problem of demodulating a unity-amplitude FM signal, let $x_k = \begin{bmatrix} \omega_k \\ \phi_k \end{bmatrix}$, $A = \begin{bmatrix} \mu_w & 0 \\ 1 & \mu_\phi \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $z_k^{(1)} = \cos(\phi_k) + v_k^{(1)}$, $z_k^{(2)} = \sin(\phi_k) + v_k^{(2)}$, where ω_k, ϕ_k, z_k and v_k denote the instantaneous frequency message, instantaneous phase, complex observations and measurement noise respectively. A zero-mean voiced speech utterance "a e i o u" was sampled at 8 kHz, for which estimates $\hat{\mu}_\omega$ = 0.97 and $\hat{\sigma}_w^2$ = 0.053 were obtained using an expectation maximization algorithm. An FM discriminator output [13],

$$z_k^{(3)} = \left(z_k^{(1)} \frac{dz_k^{(2)}}{dt} - z_k^{(2)} \frac{dz_k^{(1)}}{dt} \right)\left((z_k^{(1)})^2 + (z_k^{(1)})^2 \right)^{-1} , \quad (74)$$

serves as a benchmark and as an auxiliary frequency measurement for the above smoother. The innovations within Steps 1 and 2 are given by $\begin{bmatrix} z_k^{(1)} \\ z_k^{(2)} \\ z_k^{(3)} \end{bmatrix} - \begin{bmatrix} \cos(\hat{x}_k^{(2)}) \\ \sin(\hat{x}_k^{(2)}) \\ \hat{x}_k^{(1)} \end{bmatrix}$ and $\begin{bmatrix} \alpha_k^{(1)} \\ \alpha_k^{(2)} \\ \alpha_k^{(3)} \end{bmatrix} - \begin{bmatrix} \cos(\hat{x}_k^{(2)}) \\ \sin(\hat{x}_k^{(2)}) \\ \hat{x}_k^{(1)} \end{bmatrix}$ respectively. A unity-amplitude FM signal was synthesized using μ_ϕ = 0.99 and the SNR was varied in 1.5 dB steps from 3 dB to 15 dB. The mean-square errors were calculated over 200 realisations of Gaussian measurement noise and are shown in Fig. 9. It can be seen from the figure, that at 7.5 dB SNR, the first-order EKF improves on the FM discriminator MSE by about 12 dB. The improvement arises because the EKF

"The farther the experiment is from theory, the closer it is to the Nobel Prize." *Irène Joliot-Curie* demodulator exploits the signal model whereas the FM discriminator does not. The figure shows that the approximate minimum-variance smoother further reduces the MSE by about 2 dB, which illustrates the advantage of exploiting all the data in the time interval. In the robust designs, searches for minimum values of γ were conducted such that the corresponding Riccati difference equation solutions were positive definite over each noise realisation. It can be seen from the figure at 7.5 dB SNR that the robust EKF provides about a 1 dB performance improvement compared to the EKF, whereas the approximate minimum-variance smoother and the robust smoother performance are indistinguishable.

This nonlinear example illustrates once again that smoothers can outperform filters. Since a first-order speech model is used and the Taylor series are truncated after the first-order terms, some model uncertainty is present, and so the robust designs demonstrate a marginal improvement over the EKF.

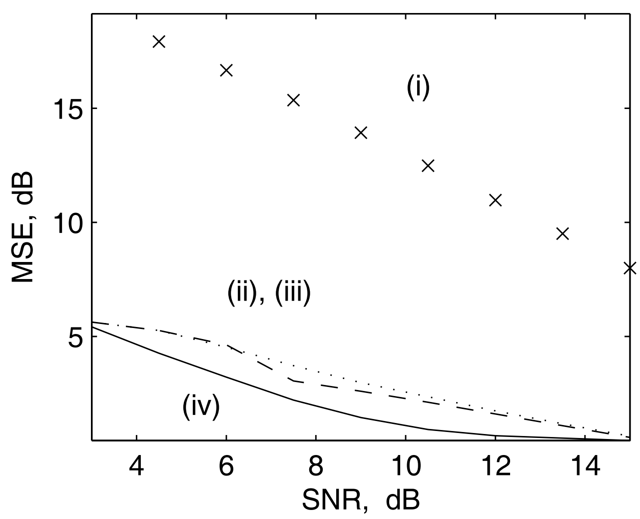


Figure 9. FM demodulation performance comparison: (i) FM discriminator (crosses), (ii) first-order EKF (dotted line), (iii) Robust EKF (dashed line), (iv) approximate minimum-variance smoother and robust smoother (solid line).

10.6 Constrained Filtering and Smoothing

10.6.1 Background

Constraints often appear within navigation problems. For example, vehicle trajectories are typically constrained by road, tunnel and bridge boundaries. Similarly, indoor pedestrian trajectories are constrained by walls and doors. However, as constraints are not easily described within state-space frameworks, many techniques for constrained filtering and smoothing are reported in the literature. An early technique for constrained filtering involves augmenting the measurement vector with perfect observations [14]. The application of the perfect-measurement approach to filtering and fixed-interval smoothing is described in [15].

"They thought I was crazy, absolutely mad." *Barbara McClintock*

Constraints can be applied to state estimates, see [16], where a positivity constraint is used within a Kalman filter and a fixed-lag smoother. Three different state equality constraint approaches, namely, maximum-probability, mean-square and projection methods are described in [17]. Under prescribed conditions, the perfect-measurement and projection approaches are equivalent [5], [18], which is identical to applying linear constraints within a form of recursive least squares.

In the state equality constrained methods [5], [16] – [18], a constrained estimate can be calculated from a Kalman filter's unconstrained estimate at each time step. Constraint information could also be embedded within nonlinear models for use with EKFs. A simpler, low-computation-cost technique that avoids EKF stablity problems and suits real-time implementation is described in [19]. In particular, an on-line procedure is proposed that involves using nonlinear functions to censor the measurements and subsequently applying the minimum-variance filter recursions. An off-line procedure for retrospective analyses is also described, where the minimum-variance fixed-interval smoother recursions are applied to the censored measurements. In contrast to the afore-mentioned techniques, which employ constraint matrices and vectors, here constraint information is represented by an exogenous input process. This approach uses the Bounded Real Lemma which enables the nonlinearities to be designed so that the filtered and smoothed estimates satisfy a performance criterion.

10.6.2 Problem Statement

The ensuing discussion concerns odd and even functions which are defined as follows. A function g_o of X is said to be odd if $g_o(-X) = -g_o(X)$. A function f_e of X is said to be even if $f_e(-X) = f_e(X)$. The product of g_o and f_e is an odd function since $g_o(-X) f_e(-X) = -g_o(X) f_e(X)$.

Problems are considered where stochastic random variables are subjected to inequality constraints. Therefore, nonlinear censoring functions are introduced whose outputs are constrained to lie within prescribed bounds. Let $\beta \in \mathbb{R}^p$ and $g_o : \mathbb{R}^p \to \mathbb{R}^p$ denote a constraint vector and an odd function of a random variable $X \in \mathbb{R}^p$ about its expected value $E\{X\}$, respectively. Define the censoring function

$$g(X) = E\{X\} + g_o(X, \beta), \tag{75}$$

where

$$g_o(X,\beta) = \begin{cases} \beta & \text{if } \beta \le X - E\{X\} \\ X - E\{X\} & \text{if } -\beta < X - E\{X\} < \beta \\ -\beta & \text{if } X - E\{X\} \le -\beta \end{cases}. \tag{76}$$

"If at first, the idea is not absurd, then there is no hope for it." *Albert Einstein*

By inspection of (75) – (76), $g(X)$ is constrained within $E\{X\} \pm \beta$. Suppose that the probability density function of X about $E\{X\}$ is even, that is, is symmetric about $E\{X\}$. Under these conditions, the expected value of $g(X)$ is given by

$$\begin{aligned} E\{g(X)\} &= \int_{-\infty}^{\infty} g(x) f_e(x)dx \\ &= E\{X\}\int_{-\infty}^{\infty} f_e(x)dx + \int_{-\infty}^{\infty} g_o(x,\beta) f_e(x)dx \\ &= E\{X\}\ . \end{aligned} \tag{77}$$

since $\int_{-\infty}^{\infty} f_e(x)dx$ = 1 and the product $g_o(x,\beta) f_e(x)$ is odd.

Thus, a constraining process can be modelled by a nonlinear function. Equation (77) states that $g(X)$ is unbiased, provided that $g_o(X,\beta)$ and $f_X(X)$ are odd and even functions about $E\{X\}$, respectively. In the analysis and examples that follow, attention is confined to systems having zero-mean inputs, states and outputs, in which case the censoring functions are also centred on zero, that is, $E\{X\} = 0$.

Let w_k = $\begin{bmatrix} w_{1,k} & \dots & w_{m,k} \end{bmatrix}^T$ ∈ $\mathbb{R}^m$ represent a stochastic white input process having an even probability density function, with $E\{w_k\} = 0$, $E\{w_j w_k^T\} = Q_k \delta_{jk}$, in which δ_{jk} denotes the Kronecker delta function. Suppose that the states of a system $\mathcal{G}: \mathbb{R}^m \to \mathbb{R}^p$ are realised by

$$x_{k+1} = A_k x_k + B_k w_k\ , \tag{78}$$

where $A_k \in \mathbb{R}^{n\times n}$ and $B_k \in \mathbb{R}^{n\times m}$. Since w_k is zero-mean, it follows that linear combinations of the states are also zero-mean. Suppose also that the system outputs, y_k, are generated by

$$y_k = \begin{bmatrix} y_{1,k} \\ \vdots \\ y_{p,k} \end{bmatrix} = \begin{bmatrix} g_o(C_{1,k}x_k, \theta_{1,k}) \\ \vdots \\ g_o(C_{p,k}x_k, \theta_{p,k}) \end{bmatrix}, \tag{79}$$

where $C_{j,k}$ is the j^{th} row of $C_k \in \mathbb{R}^{p\times m}$, θ_k = $[\theta_{1,k} \ \dots \ \theta_{p,k}]^T$ ∈ $\mathbb{R}^p$ is an input constraint process and $g_o(C_{j,k}x_k, \theta_{j,k})$, j = 1, … p, is an odd censoring function centred on zero. The outputs $y_{j,k}$ are constrained to lie within $\pm\theta_{j,k}$, that is,

$$-\theta_{j,k} \le y_{j,k} \le \theta_{j,k}\ . \tag{80}$$

For example, if the system outputs represent the trajectories of pedestrians within a building then the constraint process could include knowledge about wall, floor and ceiling positions.

"It was not easy for a person brought up in the ways of classical thermodynamics to come around to the idea that gain of entropy eventually is nothing more nor less than loss of information." *Gilbert Newton Lewis*

Similarly, a vehicle trajectory constraint process could include information about building and road boundaries.

Assume that observations $z_k = y_k + v_k$ are available, where $v_k \in \mathbb{R}^p$ is a stochastic, white measurement noise process having an even probability density function, with $E\{v_k\} = 0$, $E\{v_k\} = 0$, $E\{v_j v_k^T\} = R_k \delta_{j,k}$ and $E\{w_j v_k^T\} = 0$. It is convenient to define the stacked vectors $y = [y_1^T \; \dots \; y_N^T]^T$ and $\theta = [\theta_1^T \; \dots \; \theta_N^T]^T$. It follows that

$$\left\| y \right\|_2^2 \le \left\| \theta \right\|_2^2 . \tag{81}$$

Thus, the energy of the system's output is bounded from above by the energy of the constraint process.

The minimum-variance filter and smoother which produce estimates of a linear system's output, minimise the mean square error. Here, it is desired to calculate estimates that trade off minimum mean-square-error performance and achieve

$$\left\| \hat{y} \right\|_2^2 \le \left\| \theta \right\|_2^2 . \tag{82}$$

Note that (80) implies (81) but the converse is not true. Although estimates $\hat{y}_{j,k}$ of $y_{j,k}$ satisfying $-\theta_{j,k} \le \hat{y}_{j,k} \le \theta_{j,k}$ are desirable, the procedures described below only ensure that (82) is satisfied.

10.6.3 Constrained Filtering

A procedure is proposed in which a linear filter $\mathcal{H} : \mathbb{R}^p \to \mathbb{R}^p$ is used to calculate estimates $\hat{y}$ from zero-mean measurements z_k that are constrained using an odd censoring function to obtain

$$\underline{z}_k = \begin{bmatrix} \underline{z}_{1,k} \\ \vdots \\ \underline{z}_{p,k} \end{bmatrix} = \begin{bmatrix} g_o(z_{1,k}, \gamma^{-1}\theta_{1,k}) \\ \vdots \\ g_o(z_{p,k}, \gamma^{-1}\theta_{p,k}) \end{bmatrix}, \tag{83}$$

which satisfy

$$\left\| \underline{z} \right\|_2^2 \le \gamma^{-2} \left\| \theta \right\|_2^2 . \tag{84}$$

where $\underline{z} = [\underline{z}_1^T \; \dots \; \underline{z}_N^T]^T$, for a positive $\gamma \in \mathbb{R}$ to be designed. This design problem is depicted in Fig. 10.

"Man's greatest asset is the unsettled mind." *Isaac Asimov*

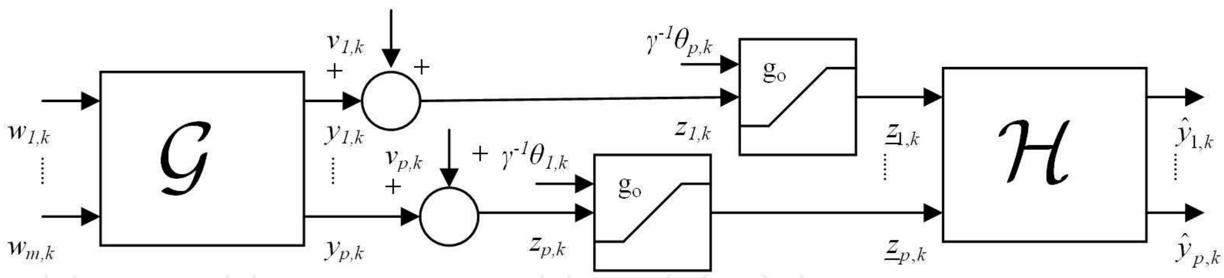


Figure 10. The constrained filtering design problem. The task is to design a scalar γ so that the outputs of a filter $\mathcal{H}$ operating on the censored zero-mean measurements $[\underline{z}_{1,k}^T \; \dots \; \underline{z}_{p,k}^T]^T$ produce output estimates $[\hat{y}_{1,k}^T \; \dots \; \hat{y}_{p,k}^T]^T$, which trade off mean square error performance and achieve $\|\hat{y}\|_2^2 \le \|\theta\|_2^2$.

Censoring the measurements is suggested as a low-implementation-cost approach to constrained filtering. Design constraints are sought for the measurement censoring functions so that the outputs of a subsequent filter satisfy the performance objective (82). The recursions akin to the minimum-variance filter are applied to calculate predicted and filtered state estimates from the constrained measurements $\underline{z}_k$ at time k. That is, the output mapping C_k is retained within the linear filter design even though nonlinearities are present with (83). The predicted states, filtered states and output estimates are respectively obtained as

$$\hat{x}_{k+1/k} = (A_k - K_k C_k)\hat{x}_{k/k-1} + K_k \underline{z}_k , \tag{85}$$

$$\hat{x}_{k/k} = (I - L_k C_k)\hat{x}_{k/k-1} + L_k \underline{z}_k , \tag{86}$$

$$\hat{y}_{k/k} = C_k \hat{x}_{k/k} , \tag{87}$$

where $L_k = P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R_k)^{-1}$, $K_k = A_k L_k$, and $P_{k/k-1} = P_{k/k-1}^T > 0$ is obtained from $P_{k/k} = P_{k/k-1} - P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R_k)^{-1} C_k P_{k/k-1}$, $P_{k+1/k} = A_k P_{k/k} A_k^T + B_k Q_k B_k^T$. Nonzero-mean sequences can be accommodated using deterministic inputs as described in Chapter 4. Since a nonlinear system output (79) and a nonlinear measurement (83) are assumed, the estimates calculated from (85) – (87) are not optimal. Some properties that are exhibited by these estimates are described below.

Lemma 3 [19]: *In respect of the filter (85) – (87) which operates on the constrained measurements (83), suppose the following:*

(i) *the probability density functions associated with w_k and v_k are even;*
(ii) *the nonlinear functions within (79) and (83) are odd; and*
(iii) *the filter is initialized with $\hat{x}_{0/0} = E\{x_0\}$.*

Then the following applies:

(i) *the predicted state estimates, $\hat{x}_{k+1/k}$, are unbiased;*
(ii) *the corrected state estimates, $\hat{x}_{k/k}$, are unbiased; and*
(iii) *the output estimates, $\hat{y}_{k/k}$, are unbiased.*

"A mind that is stretched by a new idea can never go back to its original dimensions." *Oliver Wendell Holmes*

Proof: *(i) Condition (iii) implies* $E\{\tilde{x}_{1/0}\} = 0$*, which is the initialization step of an induction argument. It follows from (85) that*

$$\hat{x}_{k+1/k} = (A_k - K_kC_k)\hat{x}_{k/k-1} + K_k(C_kx_k + v_k) + K_k(\underline{z}_k - C_kx_k - v_k)\,. \tag{88}$$

Subtracting (88) from (78) gives $\tilde{x}_{k+1/k} = (A_k - K_kC_k)\tilde{x}_{k/k-1} - B_kw_k - K_kv_k - K_k(\underline{z}_k - C_kx_k - v_k)$ *and therefore*

$$E\{\tilde{x}_{k+1/k}\} = (A_k - K_kC_k)E\{\tilde{x}_{k/k-1}\} + B_kE\{w_k\} - K_kE\{v_k\} - K_kE\{\underline{z}_k - C_kx_k - v_k\}\,. \tag{89}$$

From above assumptions, the second and third terms on the right-hand-side of (89) are zero. The property (77) implies $E\{\underline{z}_k\} = E\{z_k\} = E\{C_kx_k + v_k\}$ *and so* $E\{\underline{z}_k - C_kx_k - v_k\}$ *is zero. The first term on the right-hand-side of (89) pertains to the unconstrained Kalman filter and is zero by induction. Thus,* $E\{\tilde{x}_{k+1/k}\} = 0$.

(ii) Condition (iii) again serves as an induction assumption. It follows from (86) that

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(C_kx_k + v_k - C_k\hat{x}_{k/k-1}) + L_k(\underline{z}_k - C_kx_k - v_k)\,. \tag{90}$$

Substituting $x_k = A_{k-1}x_{k-1} + B_{k-1}w_{k-1}$ *into (90) yields* $\tilde{x}_{k/k} = (I - L_kC_k)A_{k-1}\tilde{x}_{k-1/k-1} + (I - L_kC_k)B_{k-1}w_{k-1} - L_kv_k - L_k(\underline{z}_k - C_kx_k - v_k)$ *and* $E\{\tilde{x}_{k/k}\} = (I - L_kC_k)A_{k-1}E\{\tilde{x}_{k-1/k-1}\} = (I - L_kC_k)A_{k-1} \ldots (I - L_1C_1)A_0E\{\tilde{x}_{0/0}\}$*. Hence,* $E\{\tilde{x}_{k/k}\} = 0$ *by induction.*

(iii) Defining $\tilde{y}_{k/k} = y_k - \hat{y}_{k/k} = y_k + C_k(x_k - \hat{x}_{k/k}) - C_kx_k = C_k\tilde{x}_{k/k} + y_k - C_kx_k$ *and using (77) leads to* $E\{\tilde{y}_{k/k}\} = C_kE\{\tilde{x}_{k/k}\} + E\{y_k - C_kx_k\} = C_kE\{\tilde{x}_{k/k}\} = 0$ *under condition (iii).* □

Recall that the Bounded Real Lemma (see Lemma 7 of Chapter 9) specifies a bound for a ratio of a system's output and input energies. This lemma is used to find a design for γ within (83) as described below.

Lemma 4 [19]: *Consider the filter (85) – (87) which operates on the constrained measurements (83). Let* $\underline{A}_k = A_k - K_kC_k$*,* $\underline{B}_k = K_k$*,* $\underline{C}_k = C_k(I - L_kC_k)$ *and* $\underline{D}_k = C_kL_k$ *denote the state-space parameters of the filter. Suppose for a given* $\gamma_2 > 0$*, that a solution* $M_k = M_k^T > 0$ *exists over* $k \in [1, N]$ *for the Riccati Difference equation resulting from the application of the Bounded Real Lemma to the system* $\begin{bmatrix} \underline{A}_k & \underline{B}_k \\ \underline{C}_k & \underline{D}_k \end{bmatrix}$*. Then the design* $\gamma = \gamma_2$ *within (83) results in the performance objective (82) being satisfied.*

Proof: *For the application of the Bounded Real Lemma to the filter (85) – (87), the existence of a solution* $M_k = M_k^T > 0$ *for the associated Riccati difference equation ensures that* $\|\hat{y}\|_2^2 \le \gamma_2^2\|\underline{z}\|_2^2 - x_0^TM_0x_0 \le \gamma_2^2\|\underline{z}\|_2^2$*, which together with (84) leads to (82).* □

It is argued below that the proposed filtering procedure is asymptotically stable.

"All truth passes through three stages: First, it is ridiculed; Second, it is violently opposed; and Third, it is accepted as self-evident." *Arthur Schopenhauer*

Lemma 5 [19]: *Define the filter output estimation error as* $\tilde{y} = y - \hat{y}$. *Under the conditions of Lemma 4,* $\tilde{y} \in \ell_2$.

Proof: *It follows from* $\tilde{y} = y - \hat{y}$ *that* $\|\tilde{y}\|_2 \le \|y\|_2 + \|\hat{y}\|_2$, *which together with (10) and the result of Lemma 4 yields* $\|\tilde{y}\|_2 \le 2\|\theta\|_2$, *thus the claim follows.* □

10.6.4 Constrained Smoothing

In the sequel, it is proposed that the minimum-variance fixed-interval smoother recursions operate on the censored measurements $\underline{z}_k$ to produce output estimates $\hat{y}_{k/N}$ of y_k.

Lemma 6 [19]: *In respect of the minimum-variance smoother recursions that operate on the censored measurements* $\underline{z}_k$, *under the conditions of Lemma 3, the smoothed estimates,* $\hat{y}_{k/N}$, *are unbiased.*

The proof follows *mutatis mutandis* from the approach within the proofs of Lemma 5 of Chapter 7 and Lemma 3. An analogous result to Lemma 5 is now stated.

Lemma 7 [19]: *Define the smoother output estimation error as* $\tilde{y} = y - \hat{y}$. *Under the conditions of Lemma 3,* $\tilde{y} \in \ell_2$.

The proof follows *mutatis mutandis* from that of Lemma 5. Two illustrative examples are set out below. A GPS and inertial navigation system integration application is detailed in [19].

Example 5 [19]. Consider the saturating nonlinearity

$$g_o(X,\beta) = 2\beta\pi^{-1}\arctan\left(\pi X(2\beta)^{-1}\right). \tag{91}$$

which is a continuous approximation of (76) that satisfies $|g_o(X,\beta)| \le |\beta|$ and $\frac{dg_o(X,\beta)}{dX} = \left(1 + (\pi X)^2(2\beta)^{-2}\right)^{-1} \approx 1$ when $(\pi X)^2(2\beta)^{-2} << 1$. Data was generated from (78), (79), (91), where $A = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.9 \end{bmatrix}$, $B = C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, Gaussian, white, zero-mean processes with $Q = R = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}$. The constraint vector within (80) was chosen to be fixed, namely, $\theta_k = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$, $k \in [1, 10^5]$. The limits of the observed distribution of estimates, $\hat{y}_{k/k} = \begin{bmatrix} \hat{y}_{1,k/k} \\ \hat{y}_{2,k/k} \end{bmatrix}$, arising by

"Everything we know is only some kind of approximation, because we know that we do not know all the laws yet. Therefore, things must be learned only to be unlearned again or, more likely, to be corrected." *Richard Phillips Feynman* operating the minimum-variance filter recursions on the raw data $z_k = y_k + v_k$ are indicated by the outer black region of Fig. 11. It can be seen that the filter outputs do not satisfy the performance objective (82), which motivates the pursuit of constrained techniques. A minimum value of γ_2 = 1.24 was found for the solutions of the Riccati difference equation mentioned specified within Lemma 4 to be positive definite. The filter (85) – (87) was applied to the censored measurements $\underline{z}_k = \begin{bmatrix} \underline{z}_{1,k} \\ \underline{z}_{2,k} \end{bmatrix} = \begin{bmatrix} g_o(z_{1,k}, \gamma^{-1}\theta_{1,k}) \\ g_o(z_{2,k}, \gamma^{-1}\theta_{2,k}) \end{bmatrix}$ using (91). The limits of the observed distribution of the constrained filter estimates are indicated by the inner white region of Fig. 11. The figure shows that the constrained filter estimates satisfy (82), which illustrates Lemma 5.

Example 6 [19]. Measurements were similarly synthesized using the parameters of Example 5 to demonstrate constrained fixed-interval smoother performance. A minimum value of γ_2 = 5.6 was found for the solutions of the Riccati difference equation mentioned within Lemma 4 to be positive definite. The superimposed distributions of the unconstrained and constrained smoothers are respectively indicated by the inner and outer black regions of Fig. 12. It can be seen by inspection of the figure that the constrained smoother estimates meets (80), where as those produced by the standard smoother do not.

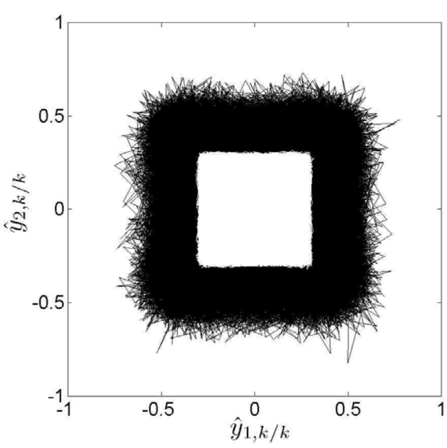


Figure 11. Superimposed distributions of filtered estimates for Example 4: unconstrained filter (outer black); and constrained filter (middle white).

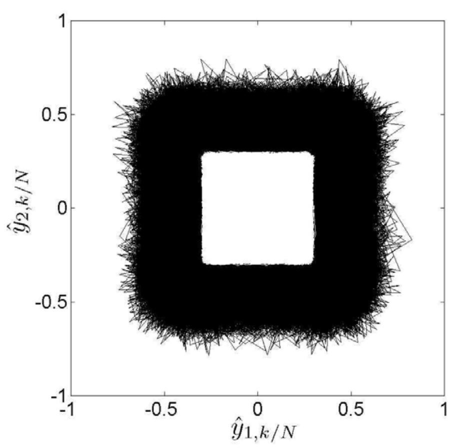


Figure 12. Superimposed distributions of smoothed estimates for Example 5: unconstrained smoother (outer black); and constrained smoother (middle white).

"An expert is a man who has made all the mistakes which can be made in a very narrow field." *Niels Henrik David Bohr*

The above examples involved searching for minimum value of γ_2 for the existence of positive definite solutions for the Riccati equation alluded to within Lemma 4. The need for a search may not be apparent as stability is guaranteed whenever a positive definite solution for the associated Riccati equation exists. Searching for a minimum γ_2 is advocated because the use of an excessively large value can lead to a nonlinearity design that is conservative and exhibits poor mean-square-error performance. If a design is still too conservative then an empirical value, namely, $\gamma_2 = \|\hat{y}\|_2 \|\underline{z}\|_2^{-1}$, may need to be considered instead.

10.7 Conclusion

In this chapter it is assumed that nonlinear systems are of the form $x_{k+1} = a_k(x_k) + b_k(w_k)$, $y_k = c_k(x_k)$, where $a_k(.)$, $b_k(.)$ and $c_k(.)$ are continuous differentiable functions. The EKF arises by linearising the model about conditional mean estimates and applying the standard filter recursions. The first, second and third-order EKFs simplified for the case of $x_k \in \mathbb{R}$ are summarised in Table 1.

The EKF attempts to produce locally optimal estimates. However, it is not necessarily stable because the solutions of the underlying Riccati equations are not guaranteed to be positive definite. The faux algebraic Riccati technique trades off approximate optimality for stability. The familiar structure of the EKF is retained but stability is achieved by selecting a positive definite solution to a faux Riccati equation for the gain design.

H_∞ techniques can be used to recast nonlinear filtering applications into a model uncertainty problem. It is demonstrated with the aid of an example that a robust EKF can reduce the mean square error when the problem is sufficiently nonlinear.

Linearised models may be applied within the previously-described smoothers in the pursuit of performance improvement. Nonlinear versions of the fixed-lag, Fraser-Potter and Rauch-Tung-Striebel smoothers are easier to implement as they are less complex. However, the application of the minimum-variance smoother can yield approximately optimal estimates when the problem becomes linear, provided that the underlying assumptions are correct. A smoother that is robust to input uncertainty is obtained by replacing the approximate error covariance correction with an H_∞ version. The resulting robust nonlinear smoother can exhibit performance benefits when uncertainty is present.

In some applications, it may be possible to censor a system's inputs, states or outputs, rather than proceed with an EKF design. It has been shown that the use of a nonlinear censoring function to constrain input measurements leads to bounded filter and smoother estimation errors.

"Most of what I learned as an entrepreneur was by trial and error." *Gordon Earl Moore*

	Linearisation	Predictor-Corrector Form
1st-order EKF	$A_k = \left.\frac{\partial a_k(x)}{\partial x}\right\|_{x=\hat{x}_{k/k}}$ $C_k = \left.\frac{\partial c_k(x)}{\partial x}\right\|_{x=\hat{x}_{k/k-1}}$ $B_k = b_k(\hat{x}_{k/k})$	$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - c_k(\hat{x}_{k/k-1}))$ $\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k})$
2nd-order EKF	$A_k = \left.\frac{\partial a_k(x)}{\partial x}\right\|_{x=\hat{x}_{k/k}}$ $C_k = \left.\frac{\partial c_k(x)}{\partial x}\right\|_{x=\hat{x}_{k/k-1}}$ $B_k = b_k(\hat{x}_{k/k})$	$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k\left(z_k - c_k(\hat{x}_{k/k-1}) - \frac{1}{2}\left.\frac{\partial^2 c_k}{\partial x^2}\right\|_{x=\hat{x}_{k/k-1}}\right)$ $\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k}) + \frac{1}{2}P_{k/k}\left.\frac{\partial^2 a_k}{\partial x^2}\right\|_{x=\hat{x}_{k/k}}$
3rd-order EKF	$A_k = \left.\frac{\partial a_k(x)}{\partial x}\right\|_{x=\hat{x}_{k/k}} + \frac{1}{6}P_{k/k}\left.\frac{\partial^2 a_k}{\partial x^2}\right\|_{x=\hat{x}_{k/k}}$ $C_k = \left.\frac{\partial c_k(x)}{\partial x}\right\|_{x=\hat{x}_{k/k-1}} + \frac{1}{6}P_{k/k-1}\left.\frac{\partial^2 c_k}{\partial x^2}\right\|_{x=\hat{x}_{k/k-1}}$ $B_k = b_k(\hat{x}_{k/k})$	$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k\left(z_k - c_k(\hat{x}_{k/k-1}) - \frac{1}{2}\left.\frac{\partial^2 c_k}{\partial x^2}\right\|_{x=\hat{x}_{k/k-1}}\right)$ $\hat{x}_{k+1/k} = a_k(\hat{x}_{k/k}) + \frac{1}{2}P_{k/k}\left.\frac{\partial^2 a_k}{\partial x^2}\right\|_{x=\hat{x}_{k/k}}$

Table 1. Summary of first, second and third-order EKFs for the case of $x_k \in \mathbb{R}$.

10.8 Problems

Problem 1. Use the following Taylor series expansion of $f(x)$

$$f(x) = f(x_0) + \frac{1}{1!}(x - x_0)^T \nabla f(x_0) + \frac{1}{2!}(x - x_0)^T \nabla^T \nabla f(x_0)(x - x_0)$$

$$+ \frac{1}{3!}(x - x_0)^T \nabla^T \nabla (x - x_0) \nabla f(x_0)(x - x_0)$$

$$+ \frac{1}{4!}(x - x_0)^T \nabla^T \nabla (x - x_0) \nabla (x - x_0) \nabla f(x_0)(x - x_0) + \ldots,$$

"The capacity to blunder slightly is the real marvel of DNA. Without this special attribute, we would still be anaerobic bacteria and there would be no music." *Lewis Thomas* to find expressions for the coefficients α_i within the functions below.

(i) $f(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)^2$.

(ii) $f(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)^2 + \alpha_3(x - x_0)^3$.

(iii) $f(x,y) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)^2$.

$$+\alpha_3(y - y_0) + \alpha_4(y - y_0)^2 + \alpha_5(x - x_0)(y - y_0)$$

(iv) $f(x,y) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)^2 + \alpha_3(x - x_0)^3$

$$+\alpha_4(y - y_0) + \alpha_5(y - y_0)^2 + \alpha_6(y - y_0)^3$$

$$+\alpha_7(x - x_0)(y - y_0) + \alpha_8(x - x_0)^2(y - y_0)$$

$$+\alpha_9(x - x_0)(y - y_0)^2 .$$

(v) $f(x,y) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)^2 + \alpha_3(x - x_0)^3 + \alpha_4(x - x_0)^4$

$$+\alpha_5(y - y_0) + \alpha_6(y - y_0)^2 + \alpha_7(y - y_0)^3 + \alpha_8(y - y_0)^4$$

$$+\alpha_9(x - x_0)(y - y_0) + \alpha_{10}(x - x_0)^2(y - y_0)$$

$$+\alpha_{11}(x - x_0)(y - y_0)^2 + \alpha_{12}(x - x_0)^3(y - y_0)$$

$$+\alpha_{13}(x - x_0)(y - y_0)^3 + \alpha_{14}(x - x_0)^2(y - y_0)^2 .$$

Problem 2. Consider a state estimation problem, where $x_{k+1} = a_k(x_k) + B_k w_k$, $y_k = c_k(x_k)$, $z_k = y_k + v_k$, in which w_k, x_k, y_k, v_k, $a_k(.)$, B_k, $c_k(.) \in \mathbb{R}$. Derive the

(i) first-order,
(ii) second-order,
(iii) third-order and
(iv) fourth-order EKFs,

assuming the required derivatives exist.

Problem 3. Suppose that an FM signal is generated by $a_{k+1} = \mu_a a_k + w_k^{(1)}$, $\omega_{k+1} = \mu_\omega \omega_k + w_k^{(2)}$, $\phi_{k+1} = \phi_k + \omega_k$, $z_k^{(1)} = a_k \cos(\phi_k) + v_k^{(1)}$ and $z_k^{(2)} = a_k \sin(\phi_k) + v_k^{(2)}$. Write down the recursions for

(i) first-order and
(ii) second-order

EKF demodulators.

"I am quite conscious that my speculations run quite beyond the bounds of true science." *Charles Robert Darwin*

Problem 4. (Continuous-time EKF) Assume that continuous-time signals may be modelled as $\dot{x}(t)$ = a(x(t)) + w(t), y(t) = c(x(t)), z(t) = y(t) + v(t), where $E\{w(t)w^T(t)\}$ = $Q(t)$ and $E\{v(t)v^T(t)\} = R(t)$.

(i) Show that approximate state estimates can be obtained from $\dot{\hat{x}}(t) = a(\hat{x}(t)) + K(t)\big(z(t) - c(\hat{x}(t))\big)$, where $K(t) = P(t)C^T(t)R^{-1}(t)$, $\dot{P}(t) = A(t)P(t) + P(t)A^T(t) - K(t)C(t)P(t) + Q(t)$, $A(t) = \left.\frac{\partial a(x)}{\partial x}\right|_{x=x(t)}$ and $C(t) = \left.\frac{\partial c(x)}{\partial x}\right|_{x=x(t)}$.

(ii) Often signal models are described in the above continuous-time setting but sampled measurements z_k of $z(t)$ are available. Write down a hybrid continuous-discrete version of the EKF in corrector-predictor form.

Problem 5. Consider a pendulum of length ℓ that subtends an angle $\theta(t)$ with a vertical line through its pivot. The pendulum's angular acceleration and measurements of its instantaneous horizontal position (from the vertical) may be modelled as $\frac{d^2\theta(t)}{d\theta^2(t)}$ = $= -\frac{g}{\ell}\sin(\theta(t))$ + $w(t)$ and $z(t) = \ell\sin(\theta(t))$ + $v(t)$, respectively, where g is the gravitational constant, $w(t)$ and $v(t)$ are stochastic inputs.

(i) Set out the pendulum's equations of motion in a state-space form and write down the continuous-time EKF for estimating $\theta(t)$ from $v(t)$.

(ii) Use Euler's first-order integration formula to discretise the above model and then detail the corresponding discrete-time EKF.

10.9 Glossary

∇f	The gradient of a function f, which is a row-vector of partial derivatives.
$\nabla^T\nabla f$	The Hessian of a function f, which is a matrix of partial derivatives.
tr(P_k)	The trace of a matrix P_k, which is the sum of its diagonal terms.
FM	Frequency modulation.
Δ_f	The forward difference operator with $\Delta_f e_k = e_k^{(i)} - e_{k-1}^{(i)}$.

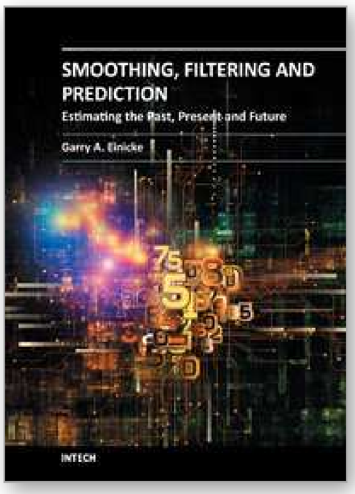
10.10 References

[1] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill Book Company, New York, 1971.
[2] A. Gelb, *Applied Optimal Estimation*, The Analytic Sciences Corporation, USA, 1974.
[3] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1979.

"What we observe is not nature itself, but nature exposed to our mode of questioning." *Werner Heisenberg*

[4] T. Söderström, *Discrete-time Stochastic Systems: Estimation and Control*, Springer-Verlag London Ltd., 2002.
[5] D. Simon, *Optimal State Estimation, Kalman H_∞ and Nonlinear Approaches*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.
[6] R. R. Bitmead, A.-C. Tsoi and P. J. Parker, "Kalman filtering approach to short time Fourier analysis", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 6, pp. 1493 – 1501, Jun. 1986.
[7] M.-A. Poubelle, R. R. Bitmead and M. Gevers, "Fake Algebraic Riccati Techniques and Stability", *IEEE Transactions on Automatic Control*, vol. 33, no. 4, pp. 379 – 381, Apr. 1988.
[8] R. R. Bitmead, M. Gevers and V. Wertz, *Adaptive Optimal Control. The thinking Man's GPC*, Prentice Hall, New York, 1990.
[9] R. R. Bitmead and Michel Gevers, "Riccati Difference and Differential Equations: Convergence, Monotonicity and Stability", In S. Bittanti, A. J. Laub and J. C. Willems (Eds.), *The Riccati Equation*, Springer Verlag, 1991.
[10] G. A. Einicke, L. B. White and R. R. Bitmead, "The Use of Fake Algebraic Riccati Equations for Co-channel Demodulation", *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2288 – 2293, Sep., 2003.
[11] C. A. Desoer and M. Vidyasagar, *Feedback Systems : Input Output Properties*, Academic Press, NewYork, 1975.
[12] G. A. Einicke and L. B. White, "Robust Extended Kalman Filtering", *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2596 – 2599, Sep., 1999.
[13] J. Aisbett, "Automatic Modulation Recognition Using Time Domain Parameters", *Signal Processing*, vol. 13, pp. 311-323, 1987.
[14] P. S. Maybeck, *Stochastic models, estimation, and control*, Academic Press, New York, vol. 1, 1979.
[15] H. E. Doran, "Constraining Kalman filter and smoothing estimates to satisfy time-varying restrictions", *Review of Economics and Statistics*, vol. 74, no. 3, pp. 568 – 572, 1992.
[16] D. Massicotte, R. Z. Morawski and A. Barwicz, "Incorporation of a Positivity Constraint Into A Kalman-Filter-Based Algorithm for Correction of Spectrometric Data", *IEEE Transactions on Instrumentation and Measurement*, vol. 44, no. 1, pp. 2 – 7, 1995.
[17] D. Simon and T. L. Chia, "Kalman Filtering with State Equality Constraints", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 128 – 136, 2002.
[18] S. J. Julier and J. J. LaViola, "On Kalman Filtering Within Nonlinear Equality Constraints", *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2774 – 2784, Jun. 2007.
[19] G. A. Einicke, G. Falco and J. T. Malos, "Bounded Constrained Filtering for GPS/INS Integration", *IEEE Transactions on Automated Control*, 2012 (to appear).

"We know nothing in reality; for truth lies in an abyss." *Democritus*



Smoothing, Filtering and Prediction - Estimating The Past, Present and Future
Edited by

ISBN 978-953-307-752-9
Hard cover, 276 pages
Publisher InTech
Published online 24, February, 2012
Published in print edition February, 2012

This book describes the classical smoothing, filtering and prediction techniques together with some more recently developed embellishments for improving performance within applications. It aims to present the subject in an accessible way, so that it can serve as a practical guide for undergraduates and newcomers to the field. The material is organised as a ten-lecture course. The foundations are laid in Chapters 1 and 2, which explain minimum-mean-square-error solution construction and asymptotic behaviour. Chapters 3 and 4 introduce continuous-time and discrete-time minimum-variance filtering. Generalisations for missing data, deterministic inputs, correlated noises, direct feedthrough terms, output estimation and equalisation are described. Chapter 5 simplifies the minimum-variance filtering results for steady-state problems. Observability, Riccati equation solution convergence, asymptotic stability and Wiener filter equivalence are discussed. Chapters 6 and 7 cover the subject of continuous-time and discrete-time smoothing. The main fixed-lag, fixed-point and fixed-interval smoother results are derived. It is shown that the minimum-variance fixed-interval smoother attains the best performance. Chapter 8 attends to parameter estimation. As the above-mentioned approaches all rely on knowledge of the underlying model parameters, maximum-likelihood techniques within expectation-maximisation algorithms for joint state and parameter estimation are described. Chapter 9 is concerned with robust techniques that accommodate uncertainties within problem specifications. An extra term within Riccati equations enables designers to trade-off average error and peak error performance. Chapter 10 rounds off the course by applying the afore-mentioned linear techniques to nonlinear estimation problems. It is demonstrated that step-wise linearisations can be used within predictors, filters and smoothers, albeit by forsaking optimal performance guarantees.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Garry Einicke (2012). Nonlinear Prediction, Filtering and Smoothing, Smoothing, Filtering and Prediction - Estimating The Past, Present and Future, (Ed.), ISBN: 978-953-307-752-9, InTech, Available from: http://www.intechopen.com/books/smoothing-filtering-and-prediction-estimating-the-past-present-and-future/nonlinear-prediction-filtering-and-smoothing

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai

www.intechopen.com

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen

IntechOpen

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen