

**ASYMPTOTIC ANALYSIS FOR
PIECEWISE LINEAR FILTERING**

Giovanni B. Di Masi
Wolfgang J. Runggaldier

June 1987
WP-87-53

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

The paper is concerned with the filtering problem for systems with piecewise linear coefficients. Such systems have recently attracted the attention of various researchers in the field of stochastic system theory.

The interest of the present study is twofold. On the one hand it provides a possible approximation to a general piecewise linear filtering problem, on the other hand it suggests the relevance of adaptive linear systems in the study of stochastic nonlinear systems.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

ABSTRACT

A discrete-time nonlinear filtering problem with piecewise linear coefficients and not necessarily Gaussian disturbances is considered. It is shown that it possesses asymptotic properties that coincide with the analogous properties of a filtering problem for a suitably randomized linear model which admits a finite-dimensional solution. The asymptotic properties are connected to the behavior of the nonlinear filters when some parameters of the distribution of the initial condition and of the signal disturbances become small. These asymptotic properties allow to consider the finite-dimensional filter as an approximate solution to the original problem. It can in fact be shown that, asymptotically, the original and the approximate models have the same conditional moments and, in particular, the same conditional mean square errors.

CONTENTS

1	Introduction	1
2	Asymptotic Analysis	4
3	Recursive Computation of the Finite-Dimensional Filter	7
	References	9

ASYMPTOTIC ANALYSIS FOR PIECEWISE LINEAR FILTERING

Giovanni B. Di Masi and Wolfgang J. Runggaldier

1. INTRODUCTION

We shall be concerned with asymptotic properties of the solution to the discrete-time nonlinear filtering problem related to the partially observable stochastic process $(x_t, y_t) \in R^n \times R^m$ described by

$$x_{t+1} = a_t(x_t) + \beta_{t+1} \quad (1.a)$$

$$y_t = c_t(x_t) + \gamma_t \quad (1.b)$$

Here a_t and c_t are continuous and piecewise linear functions, namely

$$a_t(x) = \sum_{i=1}^N [A_t(i)x + B_t(i)] I_{\pi_i}(x) \quad (2.a)$$

$$c_t(x) = \sum_{i=1}^N [C_t(i)x + D_t(i)] I_{\pi_i}(x) \quad (2.b)$$

where $(\pi_i : i = 1, \dots, N)$ is a finite partition of R^n . It is furthermore assumed that the initial condition x_0 and the disturbances β_t and γ_t are distributed according to finite mixtures of normal densities, more precisely we have (\sim means "is distributed according to" and $g(x; \mu, \sigma^2)$ denotes the normal density with mean μ and variance σ^2)

$$x_0 \sim \sum_{i=1}^{N_0} \alpha_{0i} g(x; \mu_{0i}, \sigma_{0i}^2) \quad (3.a)$$

$$\beta_t \sim \sum_{i=1}^{N_\beta} \alpha_{\beta i} g(x; \mu_{\beta i}, \sigma_{\beta i}^2) \quad (3.b)$$

$$\gamma_t \sim \sum_{i=1}^{N_\gamma} \alpha_{\gamma i} g(x; \mu_{\gamma i}, \sigma_{\gamma i}^2) \quad (3.c)$$

where $\sigma_{0i}^2, \sigma_{\beta i}^2, \sigma_{\gamma i}^2$ are positive definite. A possible representation for these random vari-

ables can be obtained in the following way. Consider a discrete random vector θ_0 which takes the values μ_{0i} , $i = 1, \dots, N_0$ with probabilities $P\{\theta_0 = \mu_{0i}\} = \alpha_{0i}$ and define the mapping $S: \{\mu_{0i}\} \rightarrow R^{n \times n}$ by

$$S(\mu_{0i}) = \sigma_{0i}$$

Assuming $v_0 \sim g(x; 0, I)$ we have that the random vector

$$x_0 = \theta_0 + S(\theta_0)v_0 \tag{4.a}$$

satisfies (3.a).

With an analogous procedure we can obtain representations for β_t and γ_t of the form

$$\beta_t = \theta_\beta + Q(\theta_\beta)v_t \tag{4.b}$$

$$\gamma_t = \theta_\gamma + R(\theta_\gamma)w_t \tag{4.c}$$

where $\{v_t: t = 0, 1, \dots\}$ and $\{w_t: t = 0, 1, \dots\}$ are independent standard Gaussian white noises, θ_β and θ_γ are discrete random variables taking values in $\{\mu_{\beta i}\}$ and $\{\mu_{\gamma i}\}$ respectively with probabilities $P(\theta_\beta = \mu_{\beta i}) = \alpha_{\beta i}$ and $P\{\theta_\gamma = \mu_{\gamma i}\} = \alpha_{\gamma i}$, and Q and R satisfy

$$Q(\mu_{\beta i}) = \sigma_{\beta i}$$

$$R(\mu_{\gamma i}) = \sigma_{\gamma i}$$

With such representations for x_0 , β_t , γ_t the mutual dependence of these random variables will be related to the joint probability $p(\theta_0, \theta_\beta, \theta_\gamma)$ and a suitable choice of the latter allows a considerable flexibility as far as the possible dependence patterns are concerned. We will therefore assume that the given nonlinear filtering problem is exactly described by equations (1), (2) with initial condition and disturbances as in (4) and with a given joint distribution $p(\theta_0, \theta_\beta, \theta_\gamma)$. It is well known [1,3] that in general such a problem admits an infinite-dimensional solution. The aim of our study here is to show that some asymptotic properties of the nonlinear filter associated with model (1), (2) coincide with those of a finite-dimensional nonlinear filter associated with a suitably chosen randomized linear model. Such filter can therefore be considered as an approximation to the original infinite-dimensional filter and for this reason the randomized linear model will be referred to as the approximate model. In order to provide a precise definition of this approximate model, we have to introduce a further random process related to the possible linear behaviors of the coefficients a_t and c_t in (2). To this end let ξ_t and η_t be the processes

defined by

$$\xi_{t+1} = a_t(\xi_t) + \theta_\beta; \xi_0 = \theta_0 \quad (5.a)$$

$$\eta_t = \sum_{i=1}^N i I_{\pi_i}(\xi_t) \quad (5.b)$$

We can now define the approximate model as:

$$x_{t+1} = A_t(\eta_t)x_t + B_t(\eta_t) + \beta_{t+1} \quad (6.a)$$

$$y_t = C_t(\eta_t)x_t + D_t(\eta_t) + \gamma_t \quad (6.b)$$

where A_t, B_t, C_t, D_t are the quantities appearing in (2) and where the initial condition x_0 and the disturbances β_t, γ_t are defined in (4). Notice that, asymptotically, η_t "tracks" the linear behaviors of a_t and c_t , namely, defining

$$\epsilon = \max \{ \sigma_{0i}, \sigma_{\beta j} : i = 1, \dots, N_0, j = 1, \dots, N_\beta \} \quad (7)$$

we have a.s.

$$\lim_{\epsilon \downarrow 0} A_t(\eta_t)x_t + B_t(\eta_t) - a_t(x_t) = 0 \quad (8.a)$$

$$\lim_{\epsilon \downarrow 0} C_t(\eta_t)x_t + D_t(\eta_t) - c_t(x_t) = 0 \quad (8.b)$$

Writing explicitly β_t and γ_t as in (4.b, c), taking into account that η_t depends only on θ_0 and θ_β , and defining $\theta = [\theta'_0, \theta'_\beta, \theta'_\gamma]'$, model (6) can be rewritten, with obvious abuse of notation, as

$$x_{t+1} = A_t(\theta)x_t + B_t(\theta) + Q(\theta)v_{t+1} \quad (9.a)$$

$$y_t = C_t(\theta)x_t + D_t(\theta) + R(\theta)w_t \quad (9.b)$$

with

$$x_0 = S_0(\theta) + S(\theta)x_0$$

where $S_0(\theta) = \theta_0$ and with a given distribution $p(\theta)$. In this case, using (8), we have a.s.

$$\lim_{\epsilon \downarrow 0} A_t(\theta)x_t + B_t(\theta) - a_t(x_t) - \theta_\beta = 0 \quad (10.a)$$

$$\lim_{\epsilon \downarrow 0} C_t(\theta)x_t + D_t(\theta) - c_t(x_t) - \theta_\gamma = 0 \quad (10.b)$$

We are now in a position to state more precisely what is the main result of this paper. We shall show, under suitable assumptions, that for any k , the k -th conditional moments of x_t given $y^t := \{y_t, \dots, y_1\}$, relative to the original model (1) and to the approximate model (9) respectively, converge to the same limit as ϵ in (7) converges to zero. This will be done in the next Section 2. In the following Section 3 we shall then show in analogy to [2] that the exact filter for the approximate model (9) is indeed finite-dimensional and derive recursive relations that allow to actually compute it.

2. ASYMPTOTIC ANALYSIS

The main result of this Section is Theorem 1 below and its Corollary 1 showing that, asymptotically for ϵ in (7) going to zero, the conditional moments of x_t given y^t , relative to the original model (1) and to the approximate model (9) respectively, converge to the same limit.

We shall need the following

ASSUMPTION A For $i = 1, \dots, N$ and all t we have

$$\lim_{\|x\| \rightarrow +\infty} \|C_t(i)x\| = +\infty \quad \square$$

Let $p_t^\epsilon(x_t, \theta)$ and $\hat{p}_t^\epsilon(x_t, \theta)$ denote the joint conditional, on y^t , distribution of (x_t, θ) for model (1) and (9) respectively. With the symbol \propto denoting proportionality and using again $g(x; \mu, \sigma^2)$ to denote the normal density, from the recursive Bayes formula we have

$$p_t^\epsilon(x_t, \theta) \propto g(y_t; c_t(x_t) + \theta_\gamma, R^2(\theta)) \cdot \int g(x_t; a_{t-1}(x_{t-1}) + \theta_\beta, Q^2(\theta)) p_{t-1}^\epsilon(x_{t-1}, \theta) dx_{t-1} \quad (11.a)$$

$$\hat{p}_t^\epsilon(x_t, \theta) \propto g(y_t; C_t(\theta)x_t + D_t(\theta), R^2(\theta)) \cdot \int g(x_t; A_{t-1}(\theta)x_{t-1} + B_{t-1}(\theta), Q^2(\theta)) \hat{p}_{t-1}^\epsilon(x_{t-1}, \theta) dx_{t-1} \quad (11.b)$$

with initial condition

$$p_0^\epsilon(x_0, \theta) = \hat{p}_0^\epsilon(x_0, \theta) = p(\theta)g(x_0; S_0(\theta), S^2(\theta)) \quad (11.c)$$

In the next Section 3 we shall derive a recursive procedure for the exact computation of $\hat{p}_t^\epsilon(x_t, \theta)$.

In the proof of Theorem 1 below we shall need the following

LEMMA 1 *Given a uniformly continuous and bounded function $f(\cdot)$, we have*

$$\lim_{\sigma \downarrow 0} \int f(x)g(x; y, \sigma^2) dx = f(y)$$

where the convergence is uniform w.r. to y .

PROOF We have for $\delta > 0$

$$\begin{aligned} & \left| \int f(x)g(x; y, \sigma^2) dx - f(y) \right| \\ & \leq \int_{|x-y| \leq \delta} |f(x) - f(y)| g(x; y, \sigma^2) dx + \int_{|x-y| > \delta} |f(x) - f(y)| g(x; y, \sigma^2) dx \end{aligned}$$

Due to the uniform continuity of $f(\cdot)$, the first integral is infinitesimal with δ for all y and σ ; due to the boundedness of $f(\cdot)$, the second integral is infinitesimal with σ for all y and δ . □

Recalling that by its definition (5a) the process ξ_t is a function of θ , for what follows, when convenient, we shall write it as $\xi_t = \xi_t(\theta)$.

We now have

THEOREM 1 *For a continuous function $f(\cdot)$ with polynomial growth we have a.s.*

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \int f(x_t) p_t^\epsilon(x_t, \theta) dx_t \\ & = \lim_{\epsilon \downarrow 0} \int f(x_t) \hat{p}_t^\epsilon(x_t, \theta) dx_t \\ & \propto K(y^t; \theta) \cdot f(\xi_t(\theta)) \end{aligned} \tag{12}$$

where

$$\begin{aligned} K(y^0, \theta) &= p(\theta) \\ K(y^t, \theta) &= K(y^{t-1}, \theta) g(y_t; c_t(\xi_t(\theta)) + \theta_\gamma, R^2(\theta)) \end{aligned} \tag{13}$$

PROOF The proof is by induction. For $t = 0$ we have

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \int f(x_0) p_0^\epsilon(x_0, \theta) dx_0 = \lim_{\epsilon \downarrow 0} \int f(x_0) \hat{p}_0^\epsilon(x_0, \theta) dx_0 \\ & = p(\theta) \lim_{\epsilon \downarrow 0} \int f(x_0) g(x_0; S_0(\theta), S^2(\theta)) dx_0 = \end{aligned}$$

$$= p(\theta)f(S_0(\theta)) = p(\theta)f(\xi_0(\theta))$$

Assume now (12) holds for $t - 1 \geq 0$. We have

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \int f(x_t) p_t^\epsilon(x_t; \theta) dx_t \\ & \propto \lim_{\epsilon \downarrow 0} \int p_{t-1}^\epsilon(x_{t-1}; \theta) \cdot \\ & \cdot \int g(y_t; c_t(x_t) + \theta_\gamma, R^2(\theta)) f(x_t) g(x_t; a_{t-1}(x_{t-1}) + \theta_\beta, Q^2(\theta)) dx_t dx_{t-1} \end{aligned}$$

By Assumption A we have that the function

$$g(y_t; c_t(x_t) + \theta_\gamma, R^2(\theta)) f(x_t) ,$$

considered as a function of x_t , is uniformly continuous and bounded. By Lemma 1 we then have that

$$\begin{aligned} & \sup_{x_t} \left| \int g(y_t; c_t(x_t) + \theta_\gamma, R^2(\theta)) f(x_t) g(x_t; a_{t-1}(x_{t-1}) + \theta_\beta, Q^2(\theta)) dx_{t-1} \right. \\ & \left. - g(y_t; c_t(a_{t-1}(x_{t-1}) + \theta_\beta) + \theta_\gamma, R^2(\theta)) f(a_{t-1}(x_{t-1}) + \theta_\beta) \right| \end{aligned}$$

is infinitesimal with ϵ . On the other hand, by the induction hypothesis

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \int p_{t-1}^\epsilon(x_{t-1}, \theta) \cdot \\ & \cdot g(y_t; c_t(a_{t-1}(x_{t-1}) + \theta_\beta) + \theta_\gamma, R^2(\theta)) \cdot f(a_{t-1}(x_{t-1}) + \theta_\beta) dx_{t-1} \\ & \propto K(y^{t-1}, \theta) g(y_t; c_t(a_{t-1}(\xi_{t-1}) + \theta_\beta) + \theta_\gamma, R^2(\theta)) \cdot \tag{14} \\ & \cdot f(a_{t-1}(\xi_{t-1}) + \theta_\beta) = K(y^t, \theta) \cdot f(\xi_t(\theta)) \end{aligned}$$

thereby accomplishing the proof of (12) as far as the original problem (1) is concerned.

In an analogous way we obtain that

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \int f(x_t) \hat{p}_t^\epsilon(x_t; \theta) dx_t \propto \\ & \propto K(y^{t-1}, \theta) \cdot g(y_t; C_t(\theta)[A_{t-1}(\theta)\xi_{t-1} + B_{t-1}(\theta)] + D_t(\theta), R^2(\theta)) \cdot \\ & \cdot f(A_{t-1}(\theta)\xi_{t-1} + B_{t-1}(\theta)) \end{aligned}$$

which by (10) and the definition of ξ_t in (5a) is equal to the right hand side of (14) thereby completing the proof. \square

Let μ_{kt}^ϵ and $\hat{\mu}_{kt}^\epsilon$ denote the conditional moments relative to the filters for the original and the approximate models respectively, namely

$$\mu_{kt}^\epsilon = \Sigma_\theta \int x^k p_t^\epsilon(x, \theta) dx \quad (15.a)$$

$$\hat{\mu}_{kt}^\epsilon = \Sigma_\theta \int x^k \hat{p}_t^\epsilon(x, \theta) dx \quad (15.b)$$

Furthermore, denote by e_t^ϵ and \hat{e}_t^ϵ respectively the conditional mean square errors of the two filters, both computed with respect to the joint conditional distribution for the original model (1), i.e.:

$$e_t^\epsilon = \Sigma_\theta \int (x_t - \mu_{1t}^\epsilon)^2 p_t^\epsilon(x_t, \theta) dx_t \quad (16.a)$$

$$\hat{e}_t^\epsilon = \Sigma_\theta \int (x_t - \hat{\mu}_{1t}^\epsilon)^2 \hat{p}_t^\epsilon(x_t, \theta) dx_t \quad (16.b)$$

We then have as an immediate consequence of Theorem 1,

COROLLARY 1 *For every t and every k*

$$\lim_{\epsilon \downarrow 0} \mu_{kt}^\epsilon = \lim_{\epsilon \downarrow 0} \hat{\mu}_{kt}^\epsilon$$

which implies, in particular, that also for the conditional mean square errors we have

$$\lim_{\epsilon \downarrow 0} e_t^\epsilon = \lim_{\epsilon \downarrow 0} \hat{e}_t^\epsilon \quad \square$$

3. RECURSIVE COMPUTATION OF THE FINITE-DEMENSIONAL FILTER

The result to be obtained in this section is given in Theorem 2 below and allows $\hat{p}_t^\epsilon(x_t, \theta)$ to be computed recursively through a finite-dimensional filter.

In what follows, given a generic nonsingular matrix M , we let

$$M^{-2} := (M^{-1})' M^{-1}$$

In analogy to [2] we now have the following

THEOREM 2 *The conditional joint distribution $\hat{p}_t^\epsilon(x_t, \theta)$ for the approximate model (9) satisfies*

$$\hat{p}_t^\epsilon(x_t, \theta) \propto [(\det S(\theta))(\det Q(\theta))^t(\det R(\theta))^t]^{-1}.$$

$$\cdot p(\theta) \left[\prod_{s=0}^t (\det N_s(\theta))^{-1/2} \right] \exp \left[-\frac{1}{2} x_t' M_t(\theta) x_t + x_t' h_t^y(\theta) + k_t^y(\theta) \right]$$

where

$$\begin{aligned} N_t(\theta) &= A_{t-1}'(\theta) Q^{-2}(\theta) A_{t-1}(\theta) + M_{t-1}(\theta); N_0(\theta) = I \\ M_t(\theta) &= C_t'(\theta) R^{-2}(\theta) C_t(\theta) + Q^{-2}(\theta) \\ &\quad - Q^{-2}(\theta) A_{t-1}(\theta) N_t^{-1}(\theta) A_{t-1}'(\theta) Q^{-2}(\theta); M_0(\theta) = S^{-2}(\theta) \\ h_t^y(\theta) &= C_t'(\theta) R^{-2}(\theta) [y_t - D_t(\theta)] + Q^{-2}(\theta) B_{t-1}(\theta) \\ &\quad + Q^{-2}(\theta) A_{t-1}(\theta) N_t^{-1}(\theta) [h_{t-1}^y(\theta) - A_{t-1}'(\theta) Q^{-2}(\theta) B_{t-1}(\theta)] ; \\ h_t^y(\theta) &= S^{-2}(\theta) S_0(\theta) \\ k_t^y(\theta) &= k_{t-1}^y(\theta) + \frac{1}{2} [h_{t-1}^y(\theta)' - B_{t-1}'(\theta) Q^{-2}(\theta) A_{t-1}(\theta)] N_t^{-1}(\theta) \cdot \\ &\quad \cdot [h_{t-1}^y(\theta) - A_{t-1}'(\theta) Q^{-2}(\theta) B_{t-1}(\theta)] - \frac{1}{2} B_{t-1}'(\theta) Q^{-2}(\theta) B_{t-1}(\theta) \\ &\quad - \frac{1}{2} [y_t - D_t(\theta)]' R^{-2}(\theta) [y_t - D_t(\theta)] ; \\ k_t^y(\theta) &= -\frac{1}{2} S_0'(\theta) S^{-2}(\theta) S_0(\theta) \end{aligned} \quad \square$$

REMARK Using the matrix equality

$$Q^{-2} - Q^{-2} A [A' Q^{-2} A + M]^{-1} A' Q^{-2} = [Q^2 + A M^{-1} A']^{-1}$$

and the positive definiteness of $S^{-2}(\theta)$ and Q^2 it is easily seen by induction that $M_t(\theta)$ and $N_t(\theta)$ are positive definite. \square

PROOF For simplicity we shall drop here the argument θ . We proceed by induction. Recalling (11.c), we immediately have the result for $t = 0$. Assuming now that the result holds for $t - 1 \geq 0$, we show it for t using the recursive formula (11.b). The induction hypothesis and a straightforward "completion of the square" lead to

$$\begin{aligned} &\int g(x_t; A_{t-1} x_{t-1} + B_{t-1}, Q^2) \hat{p}_{t-1}^\epsilon(x_{t-1}, \theta) dx_{t-1} \\ &\propto [(\det S)(\det Q)^t (\det R)^{(t-1)}]^{-1} p(\theta) \left[\prod_{s=0}^{t-1} (\det N_s)^{-1/2} \right]. \end{aligned}$$

$$\begin{aligned}
 & \cdot \exp \left\{ -\frac{1}{2} x_t' Q^{-2} x_t - \frac{1}{2} B_{t-1}' Q^{-2} B_{t-1} + x_t' Q^{-2} B_{t-1} + k_t' y_{t-1} \right\} \\
 & \cdot \exp \left\{ -\frac{1}{2} [x_t' Q^{-2} A_{t-1} - B_{t-1}' Q^{-2} A_{t-1} + h_{t-1}' y_{t-1}] [A_{t-1}' Q^{-2} A_{t-1} + M_{t-1}]^{-1} \right. \\
 & \quad \left. \cdot [A_{t-1}' Q^{-2} x_t - A_{t-1}' Q^{-2} B_{t-1} + h_{t-1}' y_{t-1}] \right\} \\
 & \cdot [\det (A_{t-1}' Q^{-2} A_{t-1} + M_{t-1})]^{-1/2}
 \end{aligned}$$

Multiplying this expression, according to (11.b), by

$$\begin{aligned}
 & g(y_t; C_t x_t + D_t, R^2) \propto (\det R)^{-1} \cdot \\
 & \cdot \exp \left[-\frac{1}{2} x_t' C_t' R^{-2} C_t + x_t' C_t' R^{-2} (y_t - D_t) - \frac{1}{2} (y_t - D_t)' R^{-2} (y_t - D_t) \right]
 \end{aligned}$$

and collecting terms, we get the desired result. \square

Theorem 2 immediately yields a finite-dimensional recursive algorithm for computing $\hat{p}_t^c(x_t, \theta)$; it consists in computing for each of the possible values of θ the recursive relations (17).

REFERENCES

- [1] Beneš, V., and I. Karatzas, "Filtering for piecewise linear drift and observation", Proc. 20th IEEE CDC Dec. 16-18 (1981), pp. 583-589.
- [2] Di Masi, G.B., and W.J. Runggaldier, "On measure transformations for combined filtering and parameter estimation in discrete time", Systems & Control Letters, 2, 1982, pp. 57-62.
- [3] Savona, C., "Filtrage linéaire par morceaux", Thèse 3ème cycle, Université de Provence Centre Saint-Charles, 1986.

Authors' permanent addresses:

Giovanni B. Di Masi, CNR-Ladseb, Corso Stati Uniti 4, I-35020 Padova, Italy

Wolfgang J. Runggaldier, Dipartimento di Matematica, Via Belzoni 7, I-35131 Padova, Italy