Journal of Applied Mathematics and Stochastic Analysis, **11**:1 (1998), 1-8.

IMPULSE CONTROL IN KALMAN-LIKE FILTERING PROBLEMS

MICHAEL V. BASIN and MARK A. PINSKY
University of Nevada at Reno, Department of Mathematics - 084
Reno, NV 89557-0045 USA
E-mail: basin@unr.edu; pinsky@unr.edu

(Received March, 1997; Revised August, 1997)

This paper develops the impulse control approach to the observation process in Kalman-like filtering problems, which is based on impulsive modeling of the transition matrix in an observation equation. The impulse control generates the jumps of the estimate variance from its current position down to zero and, as a result, enables us to obtain the filtering equations for the Kalman estimate with zero variance for all post-jump time moments. The filtering equations for the estimates with zero variances are obtained in the conventional linear filtering problem and in the case of scalar nonlinear state and nonlinear observation equations.

Key words: Kalman Filtering, Impulse Control, Stability Analysis.
AMS subject classifications: 34K, 93D.

1. Introduction

The attempts to design the optimal estimate for a dynamic system state over a certain class of observations take their origin from the fundamental Kalman and Bucy paper [3], where the optimal linear estimate was constructed. The general ideal of conditionally Gaussian filtering is known from [6]. The more complicated problem of nonlinear filtering, where only an equation for conditional density of the optimal estimate [5], but not a closed system of equations for its moments, can be obtained, was described in [1, 2]. Thus, the optimal filters were designed for various classes of discrete or continuous observations. However, none of these filters is based on application of the impulse control to the observation process, although the impulse (δ-function) approach is conventionally used for analysis of linear and nonlinear dynamic systems.

This paper presents applications of the impulse control approach to Kalman-like filtering problems. Using impulsive or pseudoimpulsive modeling of the transition matrix in an observation equation, it is possible to generate the jumps of the estimate variance from its current position down to zero and, as a result, to obtain the filtering equations for the Kalman estimate with zero variance for all post-jump time moments. The described procedure is applied to linear and scalar-state nonlinear filtering problems, where the estimates with zero variances are obtained.

The paper is organized as follows. The impulsive control is applied to the observation process and the estimates with zero variances are obtained in the conventional linear Kalman filtering problems (Section 2), in the case of a scalar nonlinear state equation (Section 3), and in the case of scalar nonlinear state and nonlinear observation equations (Section 4).

2. Impulse Control of Observations in Linear Kalman Filtering

To present the new approach and describe application of impulse control to observation process, consider first the conventional linear Kalman filtering problem [3]

$$\dot{x}(t) = Ax(t) + b(t), \quad x(t_0) = x_0, \tag{1}$$

$$y(t) = C(t)x(t) + \psi(t), \quad y(t_0) = y_0, \tag{2}$$

where $x(t) \in R^n$ is a nonobserved state and $y(t) \in R^m$ is an observation process, A is a matrix specifying asymptotic behavior of the system (1) trajectories, $C(t)$ is a transition matrix in an observation equation, $\psi(t)$ is a Gaussian noise with variance $F^{-1}(t)$, matrices $A, C(t)$ and a vector $b(t)$ are deterministic, x_0, y_0, and $\psi(t)$ are mutually independent Gaussian variables/processes.

The problem is to find the best estimate for $x(t)$ based on the observations from the initial moment t_0 to a current moment t, that is the conditional expectation $m(t) = E(x(t) \mid F_t^Y)$, where F_t^Y is the σ-algebra generated by the observations from t_0 to t.

The basic idea of the Kalman filtering is to minimize the estimate variance $P(t) = E((x(t) - m(t))(x(t) - m(t))^T \mid F_t^Y)$, where a^T denotes the transposition of a vector (or a matrix) a. However, none of the known minimizing methods is based on impulsive changes in a transition matrix $C(t)$ of an observation process, although the impulse (δ-function) approach is conventionally used for analysis of linear and nonlinear dynamic systems. This paper describes applications for the impulse control to Kalman-like filtering problems.

Let the matrix $C(t)$ change in the impulsive manner (as δ-function) at a point t^* (which may coincide with the initial point t_0), i.e.,

$$C(t) = C\delta(t - t^*), \tag{3}$$

where C is the intensity matrix of δ-function, and let the observation process (2) stop after $t = t^*$. Of course, characteristics of a physical device (specified by a matrix $C(t)$) cannot change exactly as δ-functions but may be represented as abrupt changes in $C(t)$ from its normal to peak values and back, i.e., may be described by δ-functions approximately. Moreover, in many transient stability problems for faulted systems, such pseudoimpulsive behavior of system coefficients is intentionally generated for the testing purposes in computer-aided analysis.

With the impulse transition matrix $C(t)$ given by (3), the observation equation (2) takes the form

$$y = y_0 + \int_{t_0}^{t} Cx(s)\delta(s - t^*)ds + \int_{t_0}^{t} \psi(s)\delta(s - t^*)ds,$$

and can be written as an equation of discrete-continuous observations

$$y(t) = y_0 + \int_{t_0}^{t} Cx(s)du(s) + \int_{t_0}^{t} \psi(s)du(s), \tag{4}$$

where $u(t)$ is a bounded variation function, which is equal to a Heaviside function $\chi(t - t^*)$ in this case.

The discrete-continuous observations are considered in detail in [7]. Using the results of [7], the filtering equations for the optimal estimate of the state (1) over the discrete-continuous observations (4) are written as follows

$$\begin{gathered}\dot{m}(t) = Am(t) + b(t) + P(t)C^T F(t)[\dot{y}(t) - Cm(t)\dot{u}(t)], \\ m(t_0) = m_0 = E(x_0 \mid F^Y_{t_0}), \\ \dot{P}(t) = AP(t) + P(t)A^T - P(t)C^T F(t)CP(t)\dot{u}(t), \\ P(t_0) = P_0 = E((x_0 - m_0)(x_0 - m_0)^T \mid F^Y_{t_0}).\end{gathered} \tag{5}$$

The equations (5) are equations in distributions. Their solution is defined as a vibrosolution [4], which is a function discontinuous at discontinuity points of $u(t)$. The method for determining jumps of a vibrosolution of the equations (5) at discontinuity points of $u(t)$ is given in [7].

In accordance with Theorem 2 from [7], the jumps of the best estimate $m(t)$ and its variance $P(t)$, regarded as a vibrosolution of (5), at the point t^* are computed as follows

$$\begin{gathered}\Delta m(t^*) = P(t^* -)[I + C^T F(t^*)CP(t^* -)]^{-1} C^T F(t^*)[\Delta y(t^*) - Cm(t^* -)], \\ \Delta P(t^*) = -P(t^* -)[I + C^T F(t^*)CP(t^* -)]^{-1}[C^T F(t^*)CP(t^* -)],\end{gathered} \tag{6}$$

where $\Delta f(t^*)$ is a jump of a function $f(t)$ at a point t^*, $f(t^* -)$ is a value from the left of a discontinuous function $f(t)$ at a point t^*, and I is the $n \times n$-dimensional identity matrix.

The intensity C of the impulse transition matrix can be selected in (6) in such a way that the post-jump value of $P(t)$ is approximately equal to zero up to an arbitrary precision, i.e., $0 = P(t^* +) = P(t^* -) + \lim_{\epsilon \to 0} \Delta_\epsilon P(t^*)$, where the expressions

$$\begin{gathered}\Delta_\epsilon P(t^*) = -P(t^* -)[I + \epsilon^{-1}]^{-1}\epsilon^{-1} = -P(t^* -)[\epsilon + I]^{-1}, \\ C^T F(t^*)CP(t^* -) = \epsilon^{-1},\end{gathered} \tag{7}$$

are specified by the second relation in (6), and ϵ is a positive definite matrix with a small norm. Note that the intensity matrix C can be simply selected as the matrix with the maximal possible norm, if ϵ tends to zero.

Since the Heaviside function $\chi(t - t^*) = u(t)$ in (4) is equal to zero for $t > t^*$, the rightmost terms in (5) are also vanished, and the filtering equations (5) take the following form for $t > t^*$

$$\dot{m}(t) = Am(t) + b(t), \quad m(t^*+) = m(t^*-) + \Delta m(t^* = x(t^*), \tag{8}$$

$$\dot{P}(t) = AP(t) + P(t)A^T, \quad P(t^*+) = 0, \tag{9}$$

where $m(t^*+)$ and $P(t^*+)$ are the post-jump values of the best estimate and its variance, respectively.

The homogeneous equation (9) with zero initial value has trivial zero solution. Thus, selecting the intensity C of the impulse transition matrix from (7), it is possible to obtain the filtering equations for the Kalman estimate with the best variance $P(t) = 0$ equal to zero for all post-jump time moments.

This fact seems fairly surprising. However, it can be easily understood using the following reasoning. Indeed, the impulse (δ-function) control with intensity C in the observation equation (4) perfectly compensates for the influence of the Gaussian noise ψ at a point t^*, and influence of the Gaussian noise ψ at points $t > t^*$ is ignored. Thus, since the Kalman estimate is the best estimate with respect to the σ-algebra generated by observations, the obtained estimate with zero variance is actually optimal in the class of all possible (linear and nonlinear) estimates.

Consider the stability problem for the best estimate variance $P(t)$ obtained from (6)-(9). A solution of the equation (9) with a nonzero initial value P_0 takes the form

$$P(t) = \exp(At)P_0 \exp(A^T t) \tag{10}$$

and is stable, if the matrix A has negative eigenvalues. If the latter condition is satisfied, a solution (10) will tend to zero, even if it deviates from zero due to some disturbances, for example, due to a small Gaussian noise with variance $G(t)$ in the state equation, which adds the term GG^T into the right-hand side of (9).

In the case of a few positive or critical zero/pure imaginary eigenvalues of the matrix A, the deviation of a solution (10) from zero, induced by some disturbances, will increase in view of (10). However, the variance $P(t)$ can be returned to zero by repeating application of the impulse (δ-function) control $C(t) = C_2\delta(t - t_2)$ with an appropriate intensity C_2 to the observation process (2) at a subsequent moment t_2. A moment t_2 can be selected as the moment when the deviation of $P(t)$ from zero reaches a given threshold P_2. Then, the value P_2 is substituted for $P(t^*-)$ in the equation (7), and the desired intensity C_2 of the transition matrix is determined. This procedure can be repeated at subsequent time moments $t_3, t_4, \ldots$ as many times as necessary.

Available intensities C of the transition matrix can be insufficient (i.e., too small) to return the variance $P(t)$ to zero with the desired accuracy. In this case, the described procedure should be repeated stepwise several times at close subsequent time moments, selecting the maximal possible intensity for the impulse (δ-function) control $C(t)$ at each step.

In the next sections, we consider applications of the impulse (δ-function) control to Kalman-like filtering problems in the cases of nonlinear scalar state and nonlinear observation equations.

3. Impulse Control of Observations in Kalman Filtering for a Nonlinear Scalar State Equation

Let us consider the Kalman filtering problem for a nonlinear scalar state equation ($x(t) \in R$)

$$\dot{x}(t) = f(x(t)) + b(t), \quad x(t_0) = x_0, \tag{11}$$

$$y(t) = C(t)x(t) + \psi(t), \quad y(t_0) = y_0, \tag{12}$$

where the function $f(x(t))$ is continuous, preserving the assumptions for the equations (1) and (2). The problem is to find the best estimate for $x(t)$ based on the observations from the initial moment t_0 to a current moment t, that is the conditional expectation $m(t) = E(x(t) \mid F_t^Y)$.

Since the function $f(x(t))$ is continuous, it can be approximated up to an arbitrary precision, in view of Weierstrass theorem, by a polynomial $f(x) = a_1 x + a_2 x^2 + a_3 x^3 + \ldots$, where $a_1, a_2, \ldots$ are certain coefficients. Thus, it is possible to consider an auxiliary Kalman filtering problem, which can approximate the initial one up to an arbitrary precision

$$\dot{x}(t) = a_1 x + a_2 x^2 + a_3 x^3 + \ldots + b(t), \quad x(t_0) = x_0, \tag{13}$$

$$y(t) = C(t)x(t) + \psi(t), \quad y(t_0) = y_0. \tag{14}$$

Thus, without loss of generality, the impulse control approach can be applied to the observation equation (14), which is thus transformed into the form (4). Since $m(t) = E(x(t) \mid F_t^Y)$ remains a conditionally Gaussian process in this case, the following filtering equations for the state (13) over the observations (4) can be obtained using results from [6]

$$\dot{m}(t) = a_1 m(t) + a_2 E(x^2(t) \mid F_t^Y) + a_3 E(x^3(t) \mid F_t^Y) + \ldots + b(t)$$
$$+ P(t)C^T F(t)[\dot{y}(t) - Cm(t)\dot{u}(t)], m(t_0) = m_0 = E(x_0 \mid F_{t_0}^Y), \tag{15}$$

$$\dot{P}(t) = dE[(x(t) - m(t))^2 \mid F_t^Y]/dt$$
$$- P(t)C^T F(t)CP(t)\dot{u}(t), P(t_0) = P_0 = E((x_0 - m_0)^2 \mid F_{t_0}^Y), \tag{16}$$

where $u(t) = \chi(t - t^*)$ and t^* is the point where impulse control is active.

The first term in the right-hand side of (16) can be represented, in view of (13) and (15), as follows

$$dE[(x(t) - m(t))^2 \mid F_t^Y]/dt = 2E[d(x(t) - m(t))/dt(x(t) - m(t)) \mid F_t^Y]$$
$$= 2E[a_1(x(t) - m(t))^2 + a_2(x^2(t) - E[x^2(t) \mid F_t^Y])(x(t) - m(t))$$
$$+ a_3(x^3(t) - E[x^3(t) \mid F_t^Y])(x(t) - m(t)) + \ldots \mid F_t^Y]$$
$$= 2[a_1 P + 2a_2 mP + 3a_3(P^2 + m^2 P) + \ldots],$$

because all odd central moments of a conditionally Gaussian process $m(t)$ are equal to zero [6]. Thus, the equation (16) takes the form

$$\dot{P}(t) = 2[a_1 P + 2a_2 mP + 3a_3(P^2 + m^2 P) + \ldots] - P(t)C^T F(t)CP(t)\dot{u}(t)],$$
$$P(t_0) = P_0 = E((x_0 - m_0)^2 \mid F^Y_{t_0}). \tag{17}$$

The jumps $\Delta m(t^*)$ and $\Delta P(t^*)$ at the point t^* are determined [7] by the relations (6), and the equations (15) and (17) take the following form for $t > t^*$

$$\dot{m}(t) = a_1 m(t) + a_2 E(x^2(t) \mid F^Y_t) + a_3 E(x^3(t) \mid F^Y_t) + \ldots + b(t),$$
$$m(t^* +) = m(t^* -) + \Delta m(t^*) = x(t^*), \tag{18}$$

$$\dot{P}(t) = 2[a_1 P + 2a_2 mP + 3a_3(P^2 + m^2 P) + \ldots], \quad P(t^* +) = 0. \tag{19}$$

The homogeneous equation (19) with zero initial value has trivial zero solution. Thus, selecting the intensity C of the impulse transition matrix from (7), it is possible to obtain the filtering equations for the Kalman estimate with the best variance $P(t) = 0$ equal to zero for all post-jump time moments in the case of a polynomial scalar state equation. These equations can approximate the desired filtering equations for the state (11) over the observations (12) with the transition matrix (3) up to an arbitrary precision. Based on the continuous dependence of a solution of a differential equation on its right-hand side, we conclude that the Kalman estimate for the state (11) over the observations (12) with the transition matrix (3) also has zero variance $P(t) = 0$ for all post-jump moments $t > t^*$.

Since a small deviation of an initial value P_0 of the equation (19) from zero is possible due to some disturbances (see, for example, Section 2), the local stability of the equilibrium solution $P(t) = 0$ should be assured for the convergence of a solution outgoing from $P_0 > 0$ to the equilibrium. Behavior of a solution of (19) emanating from a small neighborhood of zero, i.e., the local stability of the equilibrium $P(t) = 0$, is specified by the coefficient of the linear term $P(t)$ in the right-hand side of (19), which is equal to

$$2[a_1 + 2a_2 m + 3a_3 m^2 + \ldots + n a_n m^{n-1} + \ldots] = 2[f(m)/m].$$

Let the function $f(x)$ be such that solutions $x(t)$ of the equation (11) are bounded. Then, if the condition

$$[f(m(t))/m(t)] \le \rho < 0, \tag{20}$$

is valid for $t > t^*$, there exists a neighborhood of zero, such that a solution of (19) emanating from it will tend to the best variance $P(t) = 0$. Note that the condition (20) is equivalent to the local stability of the equilibrium solution $x(t) = 0$ of the equation $\dot{x} = f(x)$. If the condition (20) is not satisfied, a solution of (19) outgoing from $P_0 > 0$ could diverge from zero. In the latter case, the variance $P(t)$ can be returned to zero by repeating application of the impulse (δ-function) control $C(t) = C_2\delta(t - t_2)$ to the observation process (12), as it was described in Section 2.

Note that the nonlinear state equation (11) is assumed scalar for the reason of simplicity. The case of a vector nonlinear state equation should be examined additionally.

4. Impulse Control of Observations in Kalman Filtering for Nonlinear Scalar State and Nonlinear Observation Equations

Let us consider the Kalman filtering problem for nonlinear scalar state ($x(t) \in R$) and nonlinear observation equations

$$\dot{x}(t) = f(x(t)) + b(t), \quad x(t_0) = x_0, \tag{21}$$

$$y(t) = C(x(t)) + \psi(t), \quad y(t_0) = y_0, \tag{22}$$

where the function $f(x(t))$ is continuous, preserving the assumptions for the equations (1) and (2). The problem is to find the best estimate for $x(t)$ based on the observations from the initial moment t_0 to a current moment t, that is the conditional expectation $m(t) = E(x(t) \mid F_t^Y)$.

Using the polynomial approximation (13) for the state equation (21) and applying the impulse transition matrix

$$C(x(t)) = C(x(t^*))\delta(t - t^*) = Cx(t^*)\delta(t - t^*) \tag{23}$$

to the observation process (22), we obtain the following filtering problem

$$\dot{x}(t) = a_1 x + a_2 x^2 + a_3 x^3 + \ldots + b(t), \quad x(t_0) = x_0, \tag{24}$$

$$y(t) = y_0 + \int_{t_0}^{t} Cx(s)du(s) + \int_{t_0}^{t} \psi(s)du(s), \tag{25}$$

where $u(t) = \chi(t - t^*)$.

Using the theory of nonlinear filtering [1, 2], write down the equation for the best estimate in the filtering problem (24), (25), which also follows from [5]

$$\dot{m}(t) = a_1 m(t) + a_2 E(x^2(t) \mid F_t^Y) + a_3 E(x^3(t) \mid F_t^Y) + \ldots + b(t)$$

$$+ \{E[(x(t)x(t)C^T \mid F_t^Y] - m(t)E[x(t)C^T \mid F_t^Y]\}F(t)d\nu(t),$$

$$m(t_0) = m_0 = E(x_0 \mid F_{t_0}^Y), \tag{26}$$

where $d\nu(t) = dy(t) - Cm(t)d\chi(t - t^*)$. The latter term in (26) can be transformed, in view of integration with atomistic measure $d\nu(t)$ concentrated at the point t^*, as follows

$$\{E[(x(t)x(t)C^T \mid F_t^Y] - m(t)E[x(t)C^T \mid F_t^Y]\}F(t)d\nu(t)$$

$$= \{E[(x(t)x(t)) \mid F_t^Y] - m(t)m(t)\}C^T F(t)d\nu(t)$$

$$= P(t)C^T F(t)d\nu(t) = P(t)C^T F(t)[dy(t) - Cm(t)du(t)],$$

where $P(t)$ is the variance of the best estimate $m(t)$. Thus, the equation (26) takes the form

$$\dot{m}(t) = a_1 m(t) + a_2 E(x^2(t) \mid F_t^Y) + a_3 E(x^3(t) \mid F_t^Y) + \ldots + b(t)$$
$$+ P(t)C^T F(t)[dy(t) - Cm(t)du(t)], m(t_0) = m_0 = E(x_0 \mid F_{t_0}^Y). \quad (27)$$

The best estimate equation (27) depends, in this case, only on the variance $P(t)$ but on no other moments. Using the deduction from the Section 3, we obtain the following equation for the variance $P(t)$

$$\dot{P}(t) = 2[a_1 P + 2a_2 mP + 3a_3(P^2 + m^2 P) + \ldots] - P(t)C^T F(t)CP(t)\dot{u}(t),$$
$$P(t_0) = P_0 = E((x_0 - m_0)^2 \mid F_{t_0}^Y). \quad (28)$$

The jumps $\Delta m(t^*)$ and $\Delta P(t^*)$ at the point t^* are determined [7] by the relations (6), and the equations (27) and (28) respectively coincide with the equations (15) and (17) from Section 3 for $t > t^*$. Thus, selecting the intensity C of the impulse transition matrix from (7), it is possible to obtain the filtering equations for the Kalman estimate with the best variance $P(t) = 0$ equal to zero for all post-jump time moments (the equations (18) and (19) from Section 3). These equations can approximate the desired filtering equations for the state (21) over the observations (22) with the transition matrix (23) up to an arbitrary precision. Thus, we conclude that the Kalman estimate for the state (21) over the observations (22) with the impulse transition matrix (23) also has zero variance $P(t) = 0$ for all post-jump moments $t > t^*$.

The stability analysis for the estimate variance (28) at post-jump points $t > t^*$ could be done in the same way as for the estimate variance (19) from Section 3.

References

[1] Frost, P. and Kailath, T., An innovations approach to least squares estimation. Part III. Nonlinear estimation with white Gaussian noise, *IEEE Trans. Autom. Contr.* **AC-16**:6 (1971), 217-226.

[2] Fujisaki, M., Kallianpur, G. and Kunita, H., Stochastic differential equations for the nonlinear filtering problem, *Osaka J. Math.* **9** (1972), 19-40.

[3] Kalman, R.E. and Bucy, R.S., New results in linear filtering and prediction theory, *ASME Trans.* **D** (J. Basic Engineering) (1961), 95-108.

[4] Krasnosel'skii, M.A. and Pokrovskii, A.V., *Systems with Hysteresis*, Springer-Verlag, New York 1989.

[5] Kushner, H.J., On differential equations satisfied by conditional probability densities of Markov processes, *SIAM J. Control* **2** (1964), 106-119.

[6] Liptser, R.S. and Shiryayev, A.N., *Statistics of Random Processes. Part I. General Theory*, Springer-Verlag, New York 1977.

[7] Orlov, Yu.V. and Basin, M.V., On minmax filtering over discrete-continuous observations, *IEEE Trans. Autom. Contr.* **AC-40**:9 (1995), 1623-1626.