3. KOSENKO I.I., On phase trajectory construction for Hamiltonian systems in the neighbourhoods of equilibrium positions, Prikl. Mat. Mekh., 53, 4, 1989.
4. KRASNOSEL'SKII M.A., VAINIKKO G.M., ZABREIKO P.P., RUTITSKII YA.B. and STETSENKO V.YA., Approximate Solutions of Operator Equations, Nauka, Moscow, 1969.

Translated by R.L.Z.

J. Appl. Maths Mechs, Vol. 55, No. 1, pp. 29-36, 1991
Printed in Great Britain

0021-8928/91 $15.00+0.00

RECURRENT ESTIMATION AND IDENTIFICATION OF THE PARAMETERS IN NON-LINEAR DETERMINISTIC SYSTEMS*

G.N. MIL'SHTEIN and O.E. SOLOV'YEVA

Estimation of the phase states and parameters of non-linear deterministic systems of differential equations is reduced to the determination of initial data which minimize a certain functional which depends on observations and prior information. Equations are derived for an optimum non-linear filter whose realization demands repeated integration of auxiliary systems of differential equations. A modified, simpler filter, which is nearly optimum in many quite typical situations, is constructed. Consideration is given to the problem of estimation based on partly-known initial data, a special case of which is identifying the parameters of a system whose phase states are known at the initial time. In the linear case, if there is no a priori information, the results obtained here represent a deterministic version of Kalman filtering. The most constructive results in estimation have been obtained for linear systems (for general approaches see /1/, for recurrent filtration given known a priori information of a statistical nature about the initial data and noise in the object and in the observations, see /2/, for a deterministic version of recurrent estimation along game-theoretic lines, assuming known restrictions on noise, see /3/, and for a deterministic version of Kalman filtering see /4, 5/).

*1. **Statement of the problem.*** We shall consider questions relating to the estimation of non-linear systems of ordinary differential equations

$$X' = f(s, X), \quad s \geqslant t_0 \tag{1.1}$$

with observations

$$y(s) \doteq \varphi(s, X(s)), \quad s \geqslant t_0 \tag{1.2}$$

The prime denotes differentiation with respect to s, X, y are column vectors with n and m components, respectively, the approximate equality in (1.2) indicates that the observations involve an unknown degree of noise.

The identification of a parameter Λ (where Λ is an l-vector) in the system

$$X' = f(s, X, \Lambda) \tag{1.3}$$

given observations (1.2) and taking into account the relations

$$\Lambda' = 0 \tag{1.4}$$

obviously reduces to estimating the phase variables in system (1.3), (1.4) given observations (1.2) (the function φ in (1.2) may then depend on the parameter Λ: $\varphi = \varphi(s, X(s), \Lambda)$).

**Prikl.Matem.Mekhan.*, 55,1,39-47,1991

It is assumed that $f(s, x, \lambda)$ and $\varphi(s, x, \lambda)$ are such that all subsequent operations that use the continued existence in time of the solutions of the systems of differential equations and the differentiability of these solutions with respect to the initial data and parameters are admissible. These conditions are satisfied, e.g., by imposing certain smoothness conditions on f and φ and restricting the growth of f.

Define the following functional $J(X'(\cdot))$ along the solutions of system (1.1):

$$J = \alpha \,|\, X(t_0) - \bar{x} \,|^2 + \int_{t_0}^{t} g(s, X(s))\, ds, \quad \alpha \geqslant 0 \tag{1.5}$$

where $\bar{x}$ is a known vector and g is a function depending in a known way on the observations (1.2), e.g., $g(s, x) = |\, y(s) - \varphi(s, x)|^2$.

Since the solution $X(s; \vartheta, x)$ of system (1.1) is uniquely defined by the initial data

$$X(\vartheta) = x \tag{1.6}$$

for any $\vartheta \geqslant t_0$, the functional J may be treated as a function, which we denote by $J(t; \vartheta, x)$.

An estimator of the state $X(\vartheta)$ (of the parameter Λ), $\vartheta \geqslant t_0$, based on observations over the interval $[t_0, t]$ will be denoted by $x^{\wedge}_{\vartheta/t}$ $(\lambda^{\wedge}_{\vartheta/t})$. We shall try to determine it from the condition that $J(t; \vartheta, x)$ be minimized as a function of x.

Clearly,

$$J(t; s, X(s; \vartheta, x)) = J(t; \vartheta, x)$$

Hence the equality

$$x^{\wedge}_{s/t} = X(s; \vartheta, x^{\wedge}_{\vartheta/t}) \tag{1.7}$$

holds, thanks to which we may restrict ourselves to looking for an estimator $x^{\wedge}_{t_0/t}$ which is a solution of the minimization problem

$$J(t; x) \overset{\Delta}{=} J(t; t_0, x) = \alpha \,|\, x - \bar{x} \,|^2 + \int_{t_0}^{t} g(s, X(s; t_0, x))\, ds \to \min_x \tag{1.8}$$

Our immediate goal is to derive differential equations for $x^{\wedge}_{t_0/t}$.

2. ***Derivation of the equations of a non-linear filter.*** *The case* $\alpha > 0$. If $t = t_0$, the minimum in (1.8) is reached at $x = \bar{x}$. Therefore $x^{\wedge}_{t_0/t} = \bar{x}$. If $t > t_0$, we look for x in (1.8) by considering the equivalent problem of minimizing the functional (1.5) along solutions of system (1.1). This is Bolza's problem, and the necessary conditions for a solution may be obtained by using Lagrange multipliers /6/. It can be proved that the Lagrange multiplier of the integrand in (1.5) does not vanish and may therefore be equated to unity. The final form of the Euler-Lagrange equations in problem (1.5), (1.1) is

$$X' = f(s, X), \quad p' = g_x(s, X) - f_x^T(s, X) p \tag{2.1}$$

$$p(t_0) = 2\alpha(X(t_0) - \bar{x}), \quad p(t) = 0 \tag{2.2}$$

We have used the following notation: Let $y(x)$ be a scalar and $Y(x) = (y_1(x), y_2(x), \ldots, y_k(x))^T$ a column vector depending on the m variables $x = (x_1, \ldots, x_m)$. Then

$$y_x \overset{\Delta}{=} \begin{Vmatrix} \dfrac{\partial y}{\partial x_1} \\ \cdots \\ \dfrac{\partial y}{\partial x_m} \end{Vmatrix}, \quad Y_x \overset{\Delta}{=} \begin{Vmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_m} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial y_k}{\partial x_1} & \cdots & \dfrac{\partial y_k}{\partial x_m} \end{Vmatrix}$$

Clearly, a solution of the boundary-value problem (2.1), (2.2) will exist if problem (1.5), (1.1) or, what is the same, problem (1.8), is solvable. If the solution of the boundary-value problem is moreover unique, it gives an optimal solution, which is determined for each t by the initial data:

$$X(t_0) = x^{\wedge}_{t_0/t}, \quad p(t_0) = 2\alpha(x^{\wedge}_{t_0/t} - \bar{x})$$

As an abbreviation, we let $X(s; x)$, $p(s; x)$ denote the solution of system (2.1) with initial values $X(t_0) = x$, $p(t_0) = 2\alpha(x - \bar{x})$. Then the following equality holds identically with respect to t:

$$p(t; x^{\wedge}_{t_0/t}) = 0 \tag{2.3}$$

Hence we obtain the equation of the filter, whose solution is the required estimator:

$$d\hat{x}_{t_0/t}/dt = -p_x^{-1}(t; \hat{x}_{t_0/t})\, p_t(t; \hat{x}_{t_0/t}) = \\ -p_x^{-1}(t; \hat{x}_{t_0/t})\, g_x(t, X(t; \hat{x}_{t_0/t})), \quad \hat{x}_{t_0/t_0} = \bar{x} \tag{2.4}$$

The matrix $p_x(s; x)$ is part of the solution of the variational system for (2.1):

$$X_x' = f_x(s, X(s; x))X_x, \quad X_x(t_0; x) = I \tag{2.5}$$

$$p_x' = g_{xx}(s, X(s; x))X_x - \left\{\frac{\partial f_x^T}{\partial x_i}(s, X(s; x))\, p(s; x)\right\} X_x - f_x^T(s, X(s; x))\, p_x,$$

$$p_x(t_0; x) = 2\alpha I \tag{2.6}$$

The symbol I in (2.5), (2.6) denotes the $n \times n$ identity matrix, and $\left\{\frac{\partial f_x^T}{\partial x_i} p\right\}$ the matrix whose columns are $\frac{\partial f_x^T}{\partial x_i} p, \quad i = 1, 2, \ldots, n.$

Clearly, when t is close to t_0 the matrix p_x is invertible. We shall assume henceforth that p_x is invertible at all times under consideration. It follows from (2.6) that p_x^{-1} satisfies the system of equations

$$p_x^{-1\prime} = -p_x^{-1}\left[g_{xx}X_x - \left\{\frac{\partial f_x^T}{\partial x_i} p\right\} X_x\right] p_x^{-1} + p_x^{-1} f_x^T \tag{2.7}$$

$$p_x^{-1}(t_0; x) = {}^1/_2\alpha^{-1} I$$

The right-hand sides of Eq.(2.4) for the estimator $\hat{x}_{t_0/t}$ depend on $p_x^{-1}(s; x)$ and $X(s; x)$, which are functions of the $n+1$ variables s, x. These functions are part of the solution of the Cauchy problem (2.1), (2.5), (2.7) with $X(t_0) = x$, $p(t_0) = 2\alpha(x - \bar{x})$.

System (2.4) may be solved numerically by using any standard method, e.g., of the Runge-Kutta type. However, the specific form of the free terms of these equations involves repeated integration of system (2.1), (2.5) and (2.7) over increasingly long time intervals with suitable initial data.

To explain this let us take the Euler method as an example. Let $x^{\wedge(k)}$ be an approximation to $\hat{x}_{t_0/t_k}$. Then $x^{\wedge(k+1)}$ is found by solving system (2.1), (2.5), (2.7) over the interval $[t_0, t_k]$ with the initial data

$$X(t_0) = x^{\wedge(k)}, \ p(t_0) = 2\alpha(x^{\wedge(k)} - \bar{x})$$
$$X_x(t_0) = I, \ p_x^{-1}(t_0) = {}^1/_2\alpha^{-1} I$$

As a result one determines the quantities $X(t_k; x^{\wedge(k)})$, $p_x^{-1}(t_k; x^{\wedge(k)})$ and

$$x^{\wedge(k+1)} = x^{\wedge(k)} - h p_x^{-1}(t_k; x^{\wedge(k)})\, g_x(t_k, X(t_k; x^{\wedge(k)}))$$

The case $\alpha = 0$. The necessary conditions in the problem of minimizing the functional (1.5) along solutions of system (1.1) retain the form of the Euler-Lagrange Eqs.(2.1)-(2.2) even when $\alpha = 0$ (and then, of course, $p(t_0) = 0$ in (2.2)). Suppose that we have solved the boundary-value problem (2.1)-(2.2) at $\alpha = 0$ and $t = t_1 > t_0$ and have thus obtained an estimator $\hat{x}_{t_0/t_1}$. Then the behaviour of $\hat{x}_{t_0/t}$ when $t \geqslant t_1$ will be described by the filter (2.4) with initial data $\hat{x}_{t_0/t_1}$. Realization of this filter requires repeated integration of system (2.1), (2.5) and (2.6), beginning from time t_0 (we emphasize: not from time t_1). Instead of integrating system (2.6) over $[t_0, t]$ and then inverting the matrix p_x at time t (see (2.4)), we may integrate the system over $[t_0, t_1]$, invert p_x at time t_1 and then proceed to integrate system (2.7) over $[t_1, t]$ with initial data $p_x^{-1}(t_1; \hat{x}_{t_0/t})$.

Comparing the cases $\alpha = 0$ and $\alpha > 0$, we see that when $\alpha > 0$ the estimation problem is much easier to solve. This is due to the presence of the regularizing term $\alpha|x - \bar{x}|^2$, $\alpha > 0$, in the functional that represents the quality of the estimation, which in turn is conditioned by the availability of certain a pirori information about the quantities being estimated. The term $\alpha|x - \bar{x}|^2$ may sometimes be replaced by other expressions, depending on the nature of the a priori information, such as

$$\sum_{i=1}^{n} \alpha_i |x_i - \bar{x}_i|^2, \quad \alpha_i \geqslant 0$$

Clearly, such modifications of the functional (1.5) do not affect the results in any essential way.

3. ***The problem of estimation with partially known initial data.*** Let us suppose that the first k components $X_1(t_0) = \bar{x}_1, \ldots, X_k(t_0) = \bar{x}_k$ of the initial vector $X(t_0)$ are known. To estimate the remaining $n-k$ components on the basis of the observations (1.2), we consider the problem of minimizing the functional

$$J = \alpha \sum_{i=1}^{n-k} (X_{k+i}(t_0) - \bar{x}_{k+i})^2 + \int_{t_0}^{t} g(s, X(s))\, ds, \quad \alpha > 0 \tag{3.1}$$

where $\bar{x}_{k+1}$ $(i = 1, 2, \ldots, n-k)$ are known numbers. This functional depends on the last $n-k$ components of $X(t_0)$. The minimization problem (3.1), (1.1) with conditions $X_1(t_0) = \bar{x}_1, \ldots, X_k(t_0) = \bar{x}_k$ may be classified as a Bolza problem /6/ and the corresponding Euler-Lagrange equations are again of the form (2.1), but the boundary conditions are different:

$$\begin{gathered} X_i(t_0) = \bar{x}_i, \; i = 1, 2, \ldots, k, \; p_i(t_0) = 2\alpha (X_i(t_0) - \bar{x}_i) \\ i = k+1, k+2, \ldots, n; \; p_i(t) = 0, \; i = 1, 2, \ldots, n \end{gathered} \tag{3.2}$$

The solution of system (2.1) satisfying the initial conditions

$$\begin{gathered} X_i(t_0) = \bar{x}_i, \; i = 1, 2, \ldots, k, \; X_{k+i}(t_0) = x_{k+i} \triangleq \xi_i, \; i = 1, 2, \ldots, n-k \\ p_i(t_0) = p_i, \; i = 1, 2, \ldots, k, \; p_{k+i}(t_0) = 2\alpha (x_{k+i} - \bar{x}_{k+i}) \triangleq \\ 2\alpha (\xi_i - \bar{\xi}_i), \; i = 1, 2, \ldots, n-k \end{gathered}$$

is denoted by $X(t; \xi)$, $p(t; \chi) = p(t; \xi, \pi)$, where $\chi = [\xi, \pi]$, $\xi = (\xi_1, \ldots, \xi_{n-k})^T$, $\pi = (p_1, \ldots, p_k)^T$.

The analogue of Eq.(2.3) is the equation

$$p(t; \xi^\wedge, \pi^\wedge) = p(t; \chi^\wedge) = 0 \tag{3.3}$$

Clearly, $\{x^\wedge_{t_0/t}\}_i = \bar{x}_i$ $(i = 1, 2, \ldots, k)$, and $\xi^\wedge(t)$ gives the other $n-k$ coordinates of the vector $x^\wedge_{t_0/t}$. Eq.(3.3) yields a system of differential equations for $\chi^\wedge(t)$:

$$\begin{gathered} d\chi^\wedge/dt = -p_\chi^{-1}(t; \chi^\wedge)\, g_x(t, X(t; \xi^\wedge(t))), \\ \chi^\wedge(t_0) = (\bar{\xi}, 0)^T = (\bar{x}_{k+1}, \ldots, \bar{x}_n, 0, \ldots, 0)^T \end{gathered}$$

The matrix p_χ comprises two blocks: $p_\chi = [p_\xi : p_\pi]$, where p_ξ is an $n \times (n-k)$ matrix and p_π is an $(n \times k)$ matrix. The matrix p_ξ is part of the solution of the system of differential equations

$$X_\xi' = f_x(s, X(s; \xi)) X_\xi, \; X_\xi(t_0) = \begin{Vmatrix} O_{k \times (n-k)} \\ \cdots \\ I_{(n-k) \times (n-k)} \end{Vmatrix} \tag{3.4}$$

$$\begin{gathered} p_\xi' = g_{xx}(s, X(s; \xi)) X_\xi - \left\{ \frac{\partial f_x^T}{\partial x_i}(s, X(s; \xi))\, p(s; \chi) \right\} X_\xi - \\ f_x^T(s, X(s; \xi))\, p_\xi, \; p_\xi(t_0) = \begin{Vmatrix} O_{k \times (n-k)} \\ 2\alpha I_{(n-k) \times (n-k)} \end{Vmatrix} \end{gathered} \tag{3.5}$$

where O and I are the zero and identity matrices, respectively, with the appropriate number of rows (columns).

Finally, p_π satisfies the system

$$p_\pi' = -f_x^T(s, X(s; \xi))\, p_\pi, \; p_\pi(t_0) = \begin{Vmatrix} I_{k \times k} \\ O_{(n-k) \times k} \end{Vmatrix} \tag{3.6}$$

System (3.4) has less dimensions than system (2.5), while the system (3.5)-(3.6) for p_χ turns out to be separated from the start, unlike system (2.6) for p_x; all the somewhat simplifies the solution of the estimation problem.

Based on systems (3.5) and (3.6) one can also write down a system of differential equations for p_χ^{-1}.

Note that the case $\alpha = 0$ carries over in a natural way to the problem considered here.

4. *Second version of the filter equations.* In Sects.2 and 3 the equations of the filter were derived from the necessary conditions in Bolza's problem. Necessary conditions of another kind may be obtained by directly solving the minimization problem for the function $J(t; x)$ of (1.8). We have

$$J_x(t; x^\wedge_{t_0/t}) \overset{t}{\equiv} 0 \tag{4.1}$$

Hence (to fix our ideas, we are considering the case $\alpha > 0$),

$$\begin{gathered} dx^\wedge_{t_0/t}/dt = -J_{xx}^{-1}(t; x^\wedge_{t_0/t}) J_{xt}(t; x^\wedge_{t_0/t}) = \\ -J_{xx}^{-1}(t; x^\wedge_{t_0/t}) X_x^T(t; x^\wedge_{t_0/t}) g_x(t, X(t; x^\wedge_{t_0/t})), \quad x^\wedge_{t_0/t_0} = \bar{x} \end{gathered} \tag{4.2}$$

the matrix $J_{xx}(s; x)$ will be a solution of the equation

$$\begin{gathered} J'_{xx} = \left\{\frac{\partial X_x^T}{\partial x_i}(s; x) g_x(s, X(s; x))\right\} + \\ X_x^T(s; x) g_{xx}(s, X(s; x)) X_x(s; x), J_{xx}(t_0; x) = 2\alpha I \end{gathered} \tag{4.3}$$

The expression in braces in (4.3) is the matrix constructed from the columns $(\partial X_x^T/\partial x_i) g_x$, $i = 1, 2, \ldots, n$.

Because of the appearance of the second derivatives with respect to the initial data, it is generally much more difficult to realize the filter (4.2) than in the form (2.4). At the same time, there are situations in which it is more logical to employ this second version. This is precisely what happens in the problem considered in Sect.3 if the number of coordinates to be estimated is small.

Indeed, (3.1) is a function $J(t; \xi)$. As in (4.2), (4.3), we obtain

$$d\xi^\wedge/dt = -J_{\xi\xi}^{-1}(t; \xi^\wedge) X_\xi^T(t; \xi^T) g_x(t, X(t; \xi^\wedge)), \xi^\wedge(t_0) = \bar{\xi} \tag{4.4}$$

$$\begin{gathered} J'_{\xi\xi} = \left\{\frac{\partial X_\xi^T}{\partial \xi_i}(s; \xi) g_x(s, X(s; \xi))\right\} + X_\xi^T(s; \xi) g_{xx}(s, X(s; \xi)) X_\xi(s; \xi), \\ J_{\xi\xi}(t_0; \xi) = 2\alpha I_{(n-k)\times(n-k)} \end{gathered} \tag{4.5}$$

In order to realize the filters in the first version (see Sect.3) it is necessary, besides solving the systems of equations in X and X_ξ, to integrate the systems in p, p_x and $\chi^\wedge$, which contain n, n^2 and n equations, respectively, and are independent of the dimensionality of ξ; in the second version, however, we have to integrate systems for $\partial X_\xi^T/\partial \xi_i$ $(i = \overline{1, n-k})$, which all have the same number of equations $(n-k)^2 n$, the system for $J_{\xi\xi}$, which has $1/2\,(n-k)(n-k+1)$ equations, and finally the system for $\xi^\wedge$, which has $n-k$ equations. Clearly, when $n-k$ is not too large the second version is preferable.

5. *Linear filtering.* In the linear case, Eqs.(1.1), (1.2) and (1.5) have the form

$$\begin{gathered} X' = A(s)X, \ y(s) = C(s)X(s) \\ J = \alpha |X(t_0) - \bar{x}|^2 + \int_{t_0}^{t} |y(s) - C(s)X(s)|^2 ds \end{gathered} \tag{5.1}$$

where $A(s)$ is an $n \times n$ matrix and C is an $m \times n$ matrix.

In the case under consideration the filter Eqs.(4.2)-(4.3) are simplified considerably. Indeed, let $F(s, \tau)$ be a fundamental matrix of solutions of system (5.1). Then $F(s, t_0)$ is a solution of system (2.5) and $X_x^T(s; x) = F^T(s, t_0)$. Hence $\partial X_x^T/\partial x_i = 0$. The matrix $J_{xx}(s; x)$ is independent of x, because $J(s; x)$ is a quadratic function of x.

Introducing the notation $L(s) = {}^1/_2 J_{xx}(s)$, we obtain the following equations from (4.2) and (4.3):

$$\begin{gathered} dx^\wedge_{t_0/t}/dt = L^{-1}(t) F^T(t, t_0) C^T(t) (y(t) - C(t) F(t, t_0) x^\wedge_{t_0/t}), \quad x^\wedge_{t_0/t_0} = \bar{x} \\ dL^{-1}/dt = -L^{-1} F^T(t, t_0) C^T(t) C(t) F(t, t_0) L^{-1}, L^{-1}(t_0) = \alpha^{-1} I \end{gathered} \tag{5.2}$$

Since $x_t^\wedge \overset{\Delta}{=} x^\wedge_{t/t} = F(t, t_0) x^\wedge_{t_0/t}$, we use (5.2) to derive a filter which yields an estimator for the current state of the system; with the notation $M^{-1}(t) = F(t, t_0) L^{-1}(t) F^T(t, t_0)$, this equation is

$$dx_t^\wedge/dt = A(t)\,x_t^\wedge + M^{-1}(t)\,C^T(t)\,(y(t) - C(t)\,x_t^\wedge), \quad x_{t_0}^\wedge = \bar{x} \tag{5.3}$$
$$dM^{-1}/dt = A(t)\,M^{-1} + M^{-1}A^T(t) - M^{-1}C^T(t)\,C(t)\,M^{-1}, \quad M^{-1}(t_0) = \alpha^{-1}I$$

Eqs.(5.3) are well-known (see, e.g., /4, 5/) for the estimation problem with $\alpha = 0$. In this case, of course, the initial conditions at $t = t_0$, which figure in (5.3), cannot be written down. The equations may be used starting from some time $t_1 > t_0$, as already discussed at the end of Sect.2.

The filter for estimation based on partly known initial data is also simplified to a considerable degree. To obtain the equations, we introduce the following notation. We first recall that the last $n-k$ components of the vector $X(t_0)$, which have to be determined, make up a vector ξ. Denote the first k (known) components of $X(t_0)$ by $\bar{x}^{(1)} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)^T$. The matrix $F(t, t_0)$ is split into two: $F(t, t_0) = [F^{(1)}(t, t_0);\ F^{(2)}(t, t_0)]$, where $F^{(1)}$, $F^{(2)}$ are $n \times k$ and $n \times (n-k)$ matrices, respectively. With this notation the equations for $\xi^\wedge(t)$ become

$$d\xi^\wedge/dt = L_2^{-1}(t)\,[F^{(2)}(t, t_0)]^T\,C^T(t)\,(y(t) - C(t)\,F^{(1)}(t, t_0)\,\bar{x}^{(1)} - C(t)\,F^{(2)}(t, t_0)\,\xi^\wedge), \quad \xi^\wedge(t_0) = \bar{\xi} = (\bar{x}_{k+1}, \ldots, \bar{x}_n)^T \tag{5.4}$$

where $L_2^{-1}(t)$ is an $[(n-k) \times (n-k)]$ matrix which satisfies the system

$$dL_2^{-1}/dt = -L_2^{-1}\,[F^{(2)}(t, t_0)]^T\,C^T(t)\,C(t)\,F^{(2)}(t, t_0)\,L_2^{-1} \tag{5.5}$$
$$L_2^{-1}(t_0) = \alpha^{-1}I_{(n-k)\times(n-k)}$$

We also have an equation for $x_t^\wedge$:

$$dx_t^\wedge/dt = A(t)\,x_t^\wedge + F^{(2)}(t, t_0)\,L_2^{-1}(t)\,[F^{(2)}(t, t_0)]^T\,C^T(t)\,(y(t) - C(t)\,x_t^\wedge), \quad x_{t_0}^\wedge = \bar{x} \tag{5.6}$$

If $n-k$ is not too large, the filter (5.5)-(5.6) uses systems of lesser dimensions than the filter (5.3). For example, when $n-k = 1$ Eq.(5.5) is non-dimensional, while $F^{(2)}(t, t_0)$ is a vector found by solving the system determined by the first equation of (5.1).

6. Modified filter. In some fairly typical situations, the special form of Eq.(4.3) can be used to design a far simpler filter, which is close to the optimal filter (4.2).

Indeed, let $X(t; x^0)$ be the true solution. A perturbation in the observations $\delta(s) = y(s) - \varphi(s, X(s; x^0))$ is usually a mixture of high-frequency, low-amplitude oscillations. Therefore, together with $\delta(s)$, the integrals

$$\int_{t_0}^{t} \psi(s)\,\delta(s)\,ds$$

will be small independently of the length $t - t_0$ of the interval of integration, provided that $\psi(s)$ is sufficiently regular.

Let us assume for simplicity that $g = |y - \varphi|^2$. Then $g_x(s, X(s; x^0))$ has the same properties as $\delta(s)$. Assuming that from some time $x_{t_0/t}^\wedge$ which is a satisfactory estimator for x^0, we deduce that $g_x(s, X(s; x_{t_0/t}^\wedge))$ also has properties similar to those of $\delta(s)$. Noting in addition that $\partial X_x^T/\partial x_i\,(t_0; x) = 0$ $(i = 1, 2, \ldots, n)$, so that these matrices are small when t is close to t_0, we see that the filter (4.2), (4.3) may be replaced by a similar but considerably simplified modified filter:

$$dx_{t_0/t}^\vee/dt = -{}^1/_2 L^{-1}(t; x_{t_0/t}^\vee)\,X_x^T(t; x_{t_0/t}^\vee)\,g_x(t, X(t; x_{t_0/t}^\vee)), \quad x_{t_0/t_0}^\vee = \bar{x} \tag{6.1}$$

$$dL/ds = {}^1/_2 X_x^T(s; x)\,g_{xx}(s, X(s; x))\,X_x(s, x), \tag{6.2}$$
$$L(t_0; x) = \alpha I$$

It follows from (6.2) that the matrix L is positive definite, and we may therefore except the matrix $J_{xx}(t; x_{t_0/t}^\wedge)$ to be positive definite too.

A direct check shows that the function $[X_x^T(s; x)]^{-1} \cdot J_x(s; x)$ satisfies the second system of (2.1) and formula (2.2), and hence

$$p(s; x) = [X_x^T(s; x)]^{-1} J_x(s; x) \tag{6.3}$$

Hence

$$J_{xx}(s;x) = \left\{\frac{\partial X_x^T}{\partial x_i}(s;x)\,p(s;x)\right\} + X_x^T(s;x)\,p_x(s;x) \tag{6.4}$$

and by (2.3)

$$p_x(t;\hat{x}_{t_0/t}) = [X_x^T(t;\hat{x}_{t_0/t})]^{-1}\,J_{xx}(t;\hat{x}_{t_0/t}) \tag{6.5}$$

which also follows from a comparison of (2.4) and (4.2). Formula (6.5) yields a factorization of p_x as a product of a non-singular matrix and a symmetric matrix. Therefore p_x is invertible if and only if J_{xx} is invertible. Thus, the heuristic arguments given above, according to which J_{xx} should be positive definite, imply that our assumptions in Sect.2 as to p_x being a non-singular matrix were quite justifiable.

The modified filter in the first version may be obtained as follows. It should be noted that since $p(t) = 0$ the solution of the second system (2.1) can be represented by an integral in which the integrand involves g_x as a factor. Therefore (see the previous arguments) p is small and the filter constructed in Sect.2 may be modified:

$$d\check{x}_{t_0/t}/dt = -Q^{-1}(t;\check{x}_{t_0/t})\,g_x(t, X(t;\check{x}_{t_0/t})), \quad \check{x}_{t_0/t_0} = \bar{x} \tag{6.6}$$

$$\begin{gathered} Q' = g_{xx}(s, X(s;x))X_x(s;x) - f_x^T(s, X(s;x))Q \\ Q(t_0;x) = 2\alpha I \end{gathered} \tag{6.7}$$

It can be verified that

$$Q(s;x) = 2\,[X_x^T(s;x)]^{-1}L(s;x) \tag{6.8}$$

This implies that Q is non-singular and that the modified filters (6.1) and (6.6) deliver the same output. Note that if the filters (2.4) and (4.2) are constructed on the additional assumption that p_x and J_{xx} are non-singular, then the matrices Q and L in the modified filters will always be invertible.

Analogous reasoning leads to equations of a modified filter in the problem with partly known components of the initial state, in both the first and second versions.

7. ***Example.*** In connection with Mathieu's equation

$$x'' + (a^2 + 0{,}2\cos s)\,x = 0, \quad x(0) = 1, \quad x'(0) = 0 \tag{7.1}$$

with observations

$$y(s) \doteq x(s) \tag{7.2}$$

let us consider the identification problem for the parameter $\lambda = a^2$.

We introduce the notation $X_1 = x$, $X_2 = x'$ and write instead of (7.1), (7.2),

$$X_1' = X_2, \; X_1(0) = 1; \; X_2' = -(\lambda + 0{,}2\cos s)\,X_1, \; X_2(0) = 0; \; \lambda' = 0 \tag{7.3}$$

with observations

$$y(s) \doteq X_1(s) \tag{7.4}$$

The parameter λ will be sought subject to the condition that it minimize the functional

$$J = \alpha(\lambda - \bar{\lambda})^2 + \int_0^t (y(s) - X_1(s;\lambda))^2\,ds \to \min_\lambda \tag{7.5}$$

Since λ is one-dimensional, the second version of the filter is preferable. Eqs.(4.4) and (4.5) become

$$\frac{d\hat{\lambda}}{dt} = \frac{2}{J_{\lambda\lambda}(t;\hat{\lambda})}\frac{\partial X_1}{\partial\lambda}(t;\hat{\lambda})\,(y(t) - X_1(t;\hat{\lambda})), \quad \hat{\lambda}(t_0) = \bar{\lambda} \tag{7.6}$$

$$\begin{gathered} J'_{\lambda\lambda}(s;\lambda) = -2\frac{\partial^2 X_1}{\partial\lambda^2}(s;\lambda)\,(y(s) - X_1(s;\lambda)) + 2\left(\frac{\partial X_1}{\partial\lambda}(s;\lambda)\right)^2, \\ J_{\lambda\lambda}(t_0;\lambda) = 2\alpha \end{gathered} \tag{7.7}$$

In addition to these equations we must of course consider the first two equations of system (7.3) and the four equations in the first and second derivatives of X_1 and X_2 with respect to λ. Thus, realization of the filter in the second version involves integrating eight equations. In the first version, incidentally, it would have been necessary to integrate thirteen equations.

In the modified filter, we replace Eq.(7.7) by

$$L'(s;\lambda) = (\partial X_1(s;\lambda)/\partial\lambda)^2, \quad L(t_0;\lambda) = \alpha \tag{7.8}$$

The factor $2/J_{\lambda\lambda}(t;\hat{\lambda})$ in (7.6) is replaced by $1/L(t;\check{\lambda})$, and the filter is now realized by means of six equations.

Testing in a numerical experiment, we took the true value of the parameter $a^2 = \lambda$ to be $\lambda = 1$, and the noise in the observations to be $\delta(s) = \delta \sin \omega_1 s + \delta \sin \omega_2 s$. The following table lists the results of estimating λ using an optimal filter for the indicated values of $t, \bar{\lambda}, \delta, \omega_1, \omega_2$. It is evident that as t increases the estimates improve; at large t they depend only slightly on $\bar{\lambda}$ and α and are largely dependent only on δ and ω_1, ω_2. The choice of a satisfactory value of α in the functional (7.5) depends on t, δ and particularly on the proximity of $\bar{\lambda}$ to the true value of λ.

$\bar{\lambda}$	α	δ	$\omega_1 = 15,\ \omega_2 = 25$			$\omega_1 = 3,\ \omega_2 = 5$		
			$t = 0.4$	0.8	1.6	0.4	0.8	1.6
0.81	10^{-2}	0.1	0.870	1.004	0.998	0.619	0.656	1.039
		0.5	1.065	1.311	1.009	−0.201	−0.371	1.234
0.25	10^{-4}	0.1	1.703	1.121	1.003	0.092	0.467	1.024
		0.5	0.543	0.112	0.995	−10.92	−0.361	1.295

It may be observed that in this example the modified filter produces an estimate not far from that of the optimal filter.

Numerical experiments have also been carried out with random noise in the observations. The results have confirmed the high quality of the filters constructed here.

REFERENCES

1. KRASOVSKII N.N., Theory of the Control of Motion, Nauka, Moscow, 1968.
2. KALMAN R.E. and BUCY R.S., New results in linear filtering and prediction theory. Trans. ASME, Ser. D, J. Basic Engng., 83, 1, 1961.
3. KURZHANSKII A.B., Control and Observations under Conditions of Indeterminacy, Nauka, Moscow, 1977.
4. ROITENBERG YA.N., Automatic Control, Nauka, Moscow, 1978.
5. BRUMMER K. and ZIFFLING G., The Kalman-Bucy Filter, Nauka, Moscow, 1982.
6. ALEKSEYEV V.M., TIKHOMIROV V.M. and FOMIN S.V., Optimal Control, Nauka, Moscow, 1979.

Translated by D.L.