Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2015, 2(2): 60-68

Research Article ISSN: 2394 - 658X

Optimal Joint Maximum Likelihood based Estimator for Discrete Nonlinear Dynamic Systems

Ilan Rusnak

Faculty of Electrical Engineering, Technion, Haifa 32000, Israel
ilanru@rafael.co.il

ABSTRACT

The Joint Maximum Likelihood (JML) criterion is used to derive the optimal recursive-iterative estimator for discrete nonlinear dynamic systems. For linear systems this approach constitutes, in its recursive form, the structure of the Kalman Filter. The JML approach to estimation of nonlinear systems can be solved by batch formulas at the cost of extensive computational effort, i.e. with each new measurement the for derivation of the new estimate all available data has to be processed. This paper presents recursive-iterative implicit closed form solution that gives formulas of the optimal estimator, i.e. gives the value of the optimal estimated state. The computation of the estimator's gains needs the solution of non-symmetric Difference Matrix Riccati Equation (DMRE).

Key words: ML estimator, Joint maximum likelihood, non linear systems, optimal estimation, Riccati equation

INTRODUCTION

The problem of batch and recursive optimal estimation of continuous and discrete nonlinear systems has been an ongoing research area [7] before and especially since the introduction of the Kalman filter [9-10] that solved the recursive estimation problem for linear system. In course of this work the author gathered more than ninety, and counting, different approaches to estimation and smoothing of nonlinear systems, e.g. a recent one [1]. The space in this paper is too short to cover them. All these approaches are suboptimal and approximate. The exceptions are [2, 3] where for some restricted cases for nonlinear continuous systems there exist closed form solutions of the optimal filtering. These optimal solutions were derived based on Itô calculus and computation of the conditional probabilities. There are no known explicit closed form solutions of the optimal filtering for discrete nonlinear systems, i.e. solutions of the respective discrete Kolmogorov's/Fokker-Plank equation that gives the probability density function.

The most popular estimation filter of nonlinear systems is the Extended Kalman filter (EKF). The EKF filter is suboptimal. The EKF uses the Jacobians f_x and m_x of the system's differential equations functions $\dot{x} = f(x),\ y = m(x)$ for computation of the estimator's gain. The Iterated EKF (IEKF) due to Breakwell (analyzed by [14] and [6 ch. 8], [12 ch. 6]) is an EKF based algorithm that can give improved performance. The IEKF is based on heuristics rather on rigorous derivation. This iteration procedure is called in [15] Differential–Correction.

Recently the State Dependent Riccati Equation (SDRE) approach gained its place in research and application of recursive estimators of nonlinear systems [8, 11 and 13]. The SDRE approach uses the matrix *F(x)* and *M(x)* created when the system is represented in the State Dependent Coefficient (SDC) form $\dot{x} = F(x)x,\ y = M(x)x$. Such representation always exists, albeit it is not unique. The EKF, IEKF and SDRE filters are suboptimal. The conditions for the stability of the EKF are quite restrictive. The SDRE based filter needs for stability much less restrictive conditions.

The optimal discrete Kalman filter is described in details in [12]. The optimal discrete Kalman filter solution for linear systems [9,10] had been obtained as well by solving the dual of the Linear Quadratic criterion control problem [4], [5 ch. 12, 13], [6, ch. 5]. The dual of the LQ criterion is the Joint Maximum Likelihood (JML) criterion. The solution is derived by the use of the calculus of variations. This approach has been widely used for numerical implicit computations of estimators and smoothers for nonlinear dynamic systems. The optimal JML based estimator for continuous nonlinear systems has been presented in [16].

In this paper the optimal JML based estimator for discrete nonlinear systems is dealt with. This paper combines:

- the Joint Maximum Likelihood criterion;
- the State Dependent Coefficient (SDC) form representation of the discrete nonlinear system; and
- Calculus of variations to derive a recursive-iterative solution of the JML estimator for the class of nonlinear dynamic systems.

It is shown that the optimal solution includes inherently iteration, as in the IEKF, without the use of heuristics. Based on this full and exact solution suboptimal approximations can be derived. One such approximation is analyzed. It is shown that the JML optimal estimator uses a kind of a "mix" of the EKF and SDRE based estimator approaches. The state propagation equation has the form as that of the EKF. The optimal gain is computed via the solution of non-symmetric Riccati equation that uses both the Jacobians and the SDC form representation. Simulations demonstrate the performance of the optimal JML based filter. The comparison to other estimators of nonlinear systems and stability of the proposed JML filter and its approximations are beyond the scope of this paper and are the issues of ongoing research.

PROBLEM STATEMENT OF THE GENERAL PROBLEM

A general nonlinear system is dealt with. The problem statement follows closely the problem statement defined in [6, ch. 5] under the title "Statistical Methods" [6, Sect. 5.3]. Let the reality be represented by

$$\begin{aligned} &\dot{\zeta} = \varphi(\zeta, u, t), \;\; \zeta(t_o) = \zeta_o \\ &y = \mu(\zeta, t) \end{aligned} \tag{1}$$

where ζ is the real state (unknown dimension), u is the real (not necessarily known) driving force or input, y is the measured output, and the functions φ and μ represent the reality. The functions φ and μ that describe the real system are not and can't be precisely represented or are unknown precisely up to the last detail (e.g. the output measurement function may include some measurement noise or themselves exhibit random-uncertain behaviour).

For the design of the observer (or control) we use the representation-model of the reality given by

$$\begin{aligned} &x_{i+1} = f(x_i, w_i, i), \qquad x(t_o) = \bar{x}_o \\ &y_i = m(x_i, v_i, i) \end{aligned} \tag{2}$$

where x_i is the state of the model, y_i the output (the measured output), f and m are representations (model, i.e. exactly known) of the reality and thus approximation of the reality, w_i and v_i are sequences of time that represent the difference between the reality(1) and its model(1).

Problem

Derive a recursive estimator for the state of the model, x_i, from the output measurements. Here the unknowns , w_i and v_i and the initial conditions, x_o, are not random function (stochastic processes), rather can be considered as errors of unknown character [6, Sect. 5.3]. Then roughly speaking "we want to pass the solution to (2) as closely as possible, through the observations". To evaluate an estimate in the classical squares approach [6, Sect. 5.3] or the Joint Maximum Likelihood [5, ch. 13.2, problem 3] it is appropriate to minimize

$$J = \frac{1}{2}\left\{ [x_o - \hat{x}_o]^T P_o^{-1} [x_o - \hat{x}_o] + \sum_{i=0}^{k-1} w_i^T Q^{-1} w_i + \sum_{i=1}^{k} [y_i - m(x_i, v_i, i)]^T R^{-1} [y_i - m(x_i, v_i, i)] \right\} \tag{3}$$

with respect to x_i, $0 \le \iota \le k$, w_i, $0 \le \iota \le k-1$, and x_o ,subject to the model (2), where

Q is an a priori estimate of the driving force errors; R is an a priori estimate of the measurement noise errors, P_o is an a priori covariance estimate of the initial conditions errors, $\hat{x}_o$ is an a priori estimate of the initial conditions.

In other words we are looking for the representation-realization of the difference between the reality and the model, w_i, that best fits, in view of the criterion (3), the observations. The Joint Maximum Likelihood criterion is the dual of the LQ criterion for the control problem. Notice that the expectation operator does not appear in (3). This is as in contrast to the approach that results in the Kalman filter for linear stochastic systems. We do not look for solution on the average for the whole ensemble, but rather looking for the best solution with respect to criterion (3) for single sample of the modelling errors.

The rationale for this is as follows. There are problems like communication, radar, and more, where the system is known exactly (the shape of the transmitted pulse is known) and it is measured with noise and errors, and there are many replicas of this pulse and statistics of the noise is known (say Gaussian). Therefore one looks for the "best"

over the whole ensemble (doing repeated trials). However, there are cases when the system meets during its lifetime only a single sample of the noises and errors (tactical missile, airplane landing in severe weather …). This means that this system does not care about the average performance; rather it is focused on the best performance with the specific single sample. In optimization of (3) the JML resulting estimates are $w_i = \hat{w}_i$ and $x_i = \hat{x}_i$. The problem above is solvable by a batch solution that will minimize the objective (3). Here we look for a recursive solution. All vectors and matrices are of the appropriate dimensions.
Remark: If x_i, w_i, v_i are independent Gaussian random vectors, the joint probability density of (x_i, w_i, v_i) is proportional to exp(-*J*), so that minimizing *J* also corresponds to maximizing the joint probability density.

STATEMENT OF A SPECIFIC PROBLEM

The specific nonlinear system dealt with here is of the form

$$\begin{aligned} &x_{i+1} = f(x_i) + Gw_i, \qquad x_{i=0} = x_o, \\ &y_i = m(x_i) + v_i, \end{aligned} \tag{4}$$

The problem that is considered here is: given the measurements y_i, $i=1,2,\ldots k$, derive the optimal estimate of the state x_i denoted $\hat{x}_i$ by the minimization with respect to w_i and and $x_{i=0}$, of the quadratic criterion (Joint Maximum Likelihood) [5, ch. 13.2, problem 3]

$$J = \frac{1}{2}\left\{[x_o - \hat{x}_o]^T P_o^{-1}[x_o - \hat{x}_o] + \sum_{i=0}^{k-1} w_i^T Q^{-1} w_i + \sum_{i=1}^{k}[y_i - m(x_i)]^T R^{-1}[y_i - m(x_i)]\right\} \tag{5}$$

subject to the nonlinear system (4). Recursive solution is sought.

THE EXISTING RESULT FOR LINEAR

The main result for linear system is presented here. The solution is repeated here following [5, ch.13 and 12]. For the linear system

$$\begin{aligned} &x_{i+1} = Fx_i + Gw_i, \qquad\qquad x_{i=0} = x_o, \\ &y_i = Mx_i + v_i \end{aligned} . \tag{6}$$

The minimization of (5) subject to (6) leads to the well-known closed form explicit recursive formulas, the Kalman filter equations,

$$\begin{aligned} &\bar{x}_{i+1} = F\hat{x}_i, \qquad \hat{x}_{i=0} = \hat{x}_o, && \hat{x}_1(-) \\ &\hat{x}_{i+1} = \bar{x}_{i+1} + P_i M^T R^{-1}[y_{i+1} - M\bar{x}_{i+1}] && \hat{x}_1(+) \\ &N_{i+1} = FP_iF^T + GQG^T, \qquad P_{i=0} = P_o && P_1(-) \\ &P_{i+1} = \left[N_{i+1}^{-1} + M^T R^{-1} M\right]^{-1} && P_1(+) \end{aligned} \tag{7}$$

Detailed derivation of this result is in [5, chapters 12, 13]. The main consequence-conclusion of the result above is that the form of the recursive solution for linear systems (and therefore also the batch solution) for the stochastic problem (average over the ensemble); and for the statistical problem (optimization for single sample) is essentially the same. As far as the author is concerned this result is known for many years although he failed to find explicit reference to this conclusion.

THE STATE DEPENDENT COEFFICIENT FORM

The State Dependent Coefficient (SDC) form [11] of the nonlinear system (4), assuming that $f(0) = 0,\ and\ \ m(0) = 0$, is defined as

$$\begin{aligned} &x_{i+1} = F(x_i)x_i + Gw_i; \quad x_{i=0} = x_o, \\ &y_i = M(x_i)x_i + v_i \end{aligned} \tag{8}$$

This is a linear like structure. The SDC form always exists albeit it is not unique.

THE MAIN RESULT FOR NONLINEAR SYSTEM

For nonlinear system the minimization of (5) subject to (4) leads to closed form implicit, iterative and recursive formula. The optimal JML based filter for nonlinear system is

$\hat{x}_i, P_i$ from previous step

$$\begin{aligned}
&\bar{x}_{i+1} = F[\hat{x}_{i/i+1}]\hat{x}_i && \hat{x}_1(-)\\
&\hat{x}_{i/i+1} = \hat{x}_i + P_i f_x[\hat{x}_{i/i+1}]^T m_x[\hat{x}_{i+1}]^T R^{-1}\left(y_{i+1} - m[\hat{x}_{i+1}]\right)\\
&\hat{x}_{i+1} = \bar{x}_{i+1} + P_{i+1} m_x(\hat{x}_{i+1})^T R^{-1}\left[y_{i+1} - M(\hat{x}_{i+1})\bar{x}_{i+1}\right] && \hat{x}_1(+)\\
&N_{i+1} = F(\hat{x}_{i/i+1})P_i f_x(\hat{x}_{i/i+1})^T + GQG^T,\ P_{i=0} = P_o && P_1(-)\\
&P_{i+1} = \left[N_{i+1}^{-1} + m_x(\hat{x}_{i+1})^T R^{-1} M(\hat{x}_{i+1})\right]^{-1} && P_1(+)
\end{aligned} \tag{9}$$

where $f(x) =: F(x)x$, $m(x) =: M(x)x$ and $f_x(x) = \dfrac{\partial f(x)}{\partial x}$, $m_x(x) = \dfrac{\partial m(x)}{\partial x}$ are the respective Jacobians.

The optimal JML based estimator for nonlinear system is implicit. At each time propagation step there is a need to solve the set of nonlinear equations to (9) derive from $\hat{x}_i$ and y_{i+1} the new estimate $\hat{x}_{i+1}$. For linear system Eq. (9) collapses to (7) as $f_x = F, F(x) = F$ and no iteration at every time instant is needed.

The solution for discrete nonlinear systems is not as elegant as for the nonlinear continuous case [16]. To understand this let's consider first estimation step of (4) ([5, ch. 13.1] is followed closely). That is consider

$$\begin{aligned}
&x_1 = f(x_o) + Gw_o; \qquad x_o = \bar{x}_o,\\
&y_1 = m(x_1) + v_1
\end{aligned} \tag{10}$$

Let be $\hat{x}_o$ and $\bar{w}_o$ are the estimates based on information available at time *i=0*. However as y_1 is available improved estimates of x_o and w_o could be made by using the measurement y_1. In other words, measurements related to state 1 provide information about the state 0 and about the transition from state 0 to state 1, i.e. about the forcing vector w_o. These improved *estimates-smoothing estimates* are denoted by $\hat{x}_{o/1}$ and $\hat{w}_{o/1}$ to distinguish them from the estimates $\hat{x}_o$ and $\bar{w}_o$. To be least-squares estimate, the quantities $x_{o/1}$ and $w_{o/1}$ must be the values of x_o and w_o that minimize the quadratic form

$$J = \frac{1}{2}\left(x_o - \hat{x}_o\right)^T P_o^{-1}\left(x_o - \hat{x}_o\right) + \frac{1}{2}\left(w_1 - \bar{w}_o\right)^T Q^{-1}\left(w_o - \bar{w}_o\right) + \frac{1}{2}\left(y_1 - m(x_1)\right)^T R_1^{-1}\left(y_1 - m(x_1)\right) \tag{11}$$

with the constraint

$$x_1 = f(x_o) + G\bar{w}_o = F(x_o)x_o + G\bar{w}_o \tag{12}$$

where $\bar{w}_o$ is the mean of w_o. In previous section it is assumed that $w_o = 0$ for simplicity.

The exact optimal solution for one step iteration (8) is implicit and is given by

$\hat{x}_o$ from previous step (for $\bar{w}_o = 0$)

$$\begin{aligned}
&\bar{x}_1 = F(\hat{x}_{o/1})\hat{x}_o\\
&\hat{x}_{o/1} = \hat{x}_o + P_o f_x(\hat{x}_{o/1})^T m_x[\hat{x}_1]^T R^{-1}\left(y_1 - m(\hat{x}_1)\right)\\
&\hat{x}_1 = \bar{x}_1 + P_1 m_x(\hat{x}_1)^T R^{-1}\left[y_1 - M(\hat{x}_1)\bar{x}_1\right]\\
&P_1 = \left[N_1^{-1} + m_x(\hat{x}_1)^T R^{-1} M(\hat{x}_1)\right]^{-1}\\
&N_1 = F(\hat{x}_{o/1})P_o f_x(\hat{x}_{o/1})^T + GQG^T
\end{aligned} \tag{13}$$

That is at *i=1* time propagation step there is a need to find the smoothing solution for *i=0* by solving the above nonlinear set of equations (13), i.e. going backward before completing the time update. This is iteration, and with context of EKF, iterated EKF [6, Ch. 8.3, pp. 279][10] and Iterated Linear Filter-Smoother (ILFS) [6, Ch. 8.3, pp. 280], and in [11] it is called Differential-Correction. One can clearly see that as the time propagates for *i=2,3,…*, the depth of computing the smoothing backward solution deepens to derive the smoothing solution. This is not needed for the linear system case as the set of equations give explicit solution. This optimal solution reflects/possesses the properties of the batch solution, i.e. with every new measurement the estimates for the whole period should be recomputed. Thus (it seems that) the exact iterative-recursive solution for the whole period gives no computational and memory requirements advantages with respect to the batch solution.

Notice that the Iterated EKF had been derived/rationalized by heuristics [6, 12] and thus are inherently suboptimal. The presented derivation presents the optimal solution. This gives tool to systematically derive and compare iterative approximations.

SUBOPTIMAL RECURSIVE NON-ITERATIVE APPROXIMATE OF JML BASED FILTER

To simplify-approximate (9) and avoid the solution of the whole set of nonlinear equations one can approximate. There are several options of approximations. Here we chose to consider the following approximation

$$\hat{x}_{i/i+1} \cong \hat{x}_i$$

Then we have one step recursion (and no iteration) as follows

$\hat{x}_i, P_i$ from previous step

$$\begin{aligned}
&\bar{x}_{i+1} = F[\hat{x}_i]\hat{x}_i \\
&\hat{x}_{i+1} = \bar{x}_{i+1} + P_{i+1} m_x(\bar{x}_{i+1})^T R^{-1}\left[y_{i+1} - M(\bar{x}_{i+1})\bar{x}_{i+1}\right] \\
&N_{i+1} = F(\bar{x}_{i+1}) P_i f_x(\bar{x}_{i+1})^T + GQG^T, \quad P_{i=0} = P_o \\
&P_{i+1} = \left[N_{i+1}{}^{-1} + m_x(\bar{x}_{i+1})^T R^{-1} M(\bar{x}_{i+1})\right]^{-1}
\end{aligned} \tag{14}$$

This approximation gives immediately the recursive discrete Kalman filter for linear system. For nonlinear system this resembles the type of the EKF and SDDRE recursion. The gain is computed by solution of non-symmetric Riccati equation as $F(x) \neq f_x(x)$ and $M(x) \neq m_x(x)$ for non-linear system. The analogy to existing notations in KF is $\bar{x}_{i+1} = \hat{x}_1(-)$, $\hat{x}_{i+1} = \hat{x}_1(+)$, $N_{i+1} = P_1(-)$ and $P_{i+1} = P_1(+)$. The reason for the different notation is that in the nonlinear case the *P* and *N* matrices do not have the meaning of the estimation error covariance.

Remarks: The stochastic KF for linear systems gives optimal result based on minimization of the expected value of the squared error (on the ensemble), i.e. it minimizes the expectation of the square of the estimation error. It is optimal on the average. The JML based filter gives optimal result based on minimization of the squared estimation error time average per single sample of the process (kind of ergodicity?). The main consequence-conclusion of the result above is that the recursive solution for linear system for the stochastic problem (average over the ensemble); and for the statistical problem (optimization for single sample) are essentially the same (expectation over the ensemble= mean in time for single sample). This is not true in general for nonlinear systems.

COMPARISON OF EKF, SDDRE AND JOINT MAXIMUM LIKELIHOOD BASED ESTIMATORS

In this section the EKF and SDDRE approaches are presented in order to point out the differences between them and the JML approach. The nonlinear system is (4, 8)

$$\begin{aligned}
x_{i+1} &= f(x_i) + Gw_i = F(x_i)x_i + Gw_i; \quad x_{i=0} = x_o, \\
y_i &= m(x_i) + v_i = M(x_i)x_i + v_i
\end{aligned} \tag{15}$$

First for completeness of presentation the EKF and SDDRE based filters are presented.

The EKF

The EKF filter is

$\hat{x}_i, P_i$ from previous step

$$\begin{aligned}
&\bar{x}_{i+1} = f(\hat{x}_i) = F(\hat{x}_i)\hat{x}_i \\
&\hat{x}_{i+1} = \bar{x}_{i+1} + P_{i+1} m_x(\bar{x}_{i+1})^T R^{-1}\left[y_{i+1} - m(\bar{x}_{i+1})\right] \\
&N_{i+1} = f_x(\bar{x}_{i+1}) P_i f_x(\bar{x}_{i+1})^T + GQG^T, \quad P_{i=0} = P_o \\
&P_{i+1} = \left[N_{i+1}{}^{-1} + m_x(\bar{x}_{i+1})^T R^{-1} m_x(\bar{x}_{i+1})\right]^{-1}
\end{aligned} \tag{16}$$

The SDDRE Estimator

The SDC form is (8).Then the SDDRE based filter is

$\hat{x}_i, P_i$ from previous step

$$\begin{aligned}
&\bar{x}_{i+1} = f(\hat{x}_i) = F(\hat{x}_i)\hat{x}_i \\
&\hat{x}_{i+1} = \bar{x}_{i+1} + P_{i+1} M(\bar{x}_{i+1})^T R^{-1}\left[y_{i+1} - m(\bar{x}_{i+1})\right] \\
&N_{i+1} = F(\bar{x}_{i+1}) P_i F(\bar{x}_{i+1})^T + GQG^T, \quad P_{i=0} = P_o \\
&P_{i+1} = \left[N_{i+1}{}^{-1} + M(\bar{x}_{i+1})^T R^{-1} M(\bar{x}_{i+1})\right]^{-1}
\end{aligned} \tag{17}$$

Comparison of EKF, SDDRE and Joint Maximum Likelihood based Filters

From the equations above one can see:

(i) for nonlinear system with linear measurements the difference equation that propagates the estimated state has similar structure, i.e.

$$\begin{aligned} \bar{x}_{i+1} &= f(\hat{x}_i) = F(\hat{x}_i)\hat{x}_i \\ \hat{x}_{i+1} &= \bar{x}_{i+1} + P_{i+1} m_x(\bar{x}_{i+1})^T R^{-1}\left[y_{i+1} - m(\bar{x}_{i+1})\right], \; EKF \; and \; JML \\ \hat{x}_{i+1} &= \bar{x}_{i+1} + P_{i+1} M(\bar{x}_{i+1})^T R^{-1}\left[y_{i+1} - m(\bar{x}_{i+1})\right], \; SDDRE \end{aligned} \tag{18}$$

(ii) the fundamental difference between the EKF, SDDRE filter and Joint Maximum Likelihood based optimal estimator is the computation of the P matrix from which the filter gain is computed. The Riccati matrix equation associated to:

- EKF is symmetric and uses the Jacobian of the system dynamics equation $f_x(x)$;
- SDDRE filter is symmetric and uses the SDC representation of the system dynamics equation $F(x)$and $M(x)$;
- JML optimal estimator is asymmetric and uses both the Jacobian of the system equation $f_x(x)$, $m_x(x)$ and the State Dependent (SDC) form representation of the system dynamics equation $F(x)$, $M(x)$. Table 1 summarized the conclusions above.

Table 1: Comparison of Riccati Matrix Equation of EKF, SDDRE Filter and JML Filter

	$N_{i+1} = AP_iB^T + C,\; P_{i+1} = \left[N_{i+1}{}^{-1} + D^T R^{-1} E\right]^{-1}$					$PZ^T R^{-1}$
	A	B	C	D	E	Z
EKF	f_x	f_x	GQG^T	m_x	m_x	m_x
SDDRE-F	F	F	GQG^T	M	M	M
JMLF	F	f_x	GQG^T	m_x	M	m_x

EXAMPLES

The performance of the JML based filter (9) is demonstrated here. In this section it is assumed that the sampling interval is sufficiently small such that instead solving the complete iteration scheme the continuous version of the JML based filter [16] is implemented and its performance is demonstrated.

Van der Pol equation

This section demonstrates the performance of the JML based estimator on the nonlinear Van der Pol differential equation driven by band limited white noise and nonlinear noise corrupted measurement. The Van der Pol equation in matrix from is

$$\frac{d}{dt}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{2c}{m}(x^2-1) \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w \quad \text{where } w \text{ is a system driving noise} \tag{19}$$

$$y = \frac{x}{\sqrt{1+x^2}} + v$$

Then we have

$$f(x) = \begin{bmatrix} \dot{x} \\ -\frac{k}{m}x - \frac{2c}{m}(x^2-1)\dot{x} \end{bmatrix} \tag{20}$$

The SCD form system matrix is selected as

$$F(x) = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{2c}{m}(x^2-1) \end{bmatrix} \tag{21}$$

and the respective Jacobian is

$$f_x(x) = \begin{bmatrix} 0 & 1 \\ -\left(\frac{k}{m} + \frac{4c}{m}x\dot{x}\right) & -\frac{2c}{m}(x^2-1) \end{bmatrix} \tag{22}$$

And $$m(x)=\frac{x}{\sqrt{1+x^2}} \qquad M(x)=\begin{bmatrix}\frac{1}{\sqrt{1+x^2}} & 0\end{bmatrix} \qquad m_x(x)=\begin{bmatrix}\frac{1}{\left(1+x^2\right)^{3/2}} & 0\end{bmatrix} \tag{23}$$

For linear time-varying Kalman Filter for existence of solution of the Riccati equation it is necessary and sufficient that the respective observability and controllability Gramians are uniformly completely bounded [17]. (From above and from below.) The observability and controllability analysis of the presented example is important however it is beyond the scope of the paper.

The system and the JML based Filter were implemented in SIIMULINK® with the following parameters:

m =1;
c =0.1;
k =1;
R_c =1e-3[1/Hz] (spectral density of the measurement noise - v)
Q_c =1e0 [$(1/\sec^2)^2$/Hz] (spectral density of the system driving noise - w)
P0=[0 0 ; 0 0]; (initial condition of the P matrix)

The measurement noise and system driving noises are white in 100 [rad/sec] bandwidth. The following cases are considered:

(i) No actual measurement noise and no actual system driving noise
(ii) measurement noise and system driving noise

Figs 1 and 7, present the measured output, $y=x+v$, and the estimated output, $\hat{y}$, versus time. The transient as well the quality of estimation can be seen.

Figs 2 and 8 present the real position, x, and the estimated position state, $\hat{x}$, versus time. The transient as well the quality of estimation can be seen.

Figs 3and 9 present the output estimation errors and position estimation error versus time.

Figs 4 and 10 present the real velocity – $\dot{x}$ and the estimated velocity – $\hat{\dot{x}}$ versus time.

Figs 5 and 11 present the filter's gains, K1- gain of the position state, K2 - gain of the velocity state versus time.

Figs 6 and 12 present the terms of the solution of the matrix Riccati equation - P versus time. One can see that K1=P11/R and K2=P21/R. One can clearly see that the P matrix is non-symmetric P12≠P21. The difference in this example is small.

These figures demonstrate the performance of the JML based estimator.

No actual measurement noise and no actual system driving noise

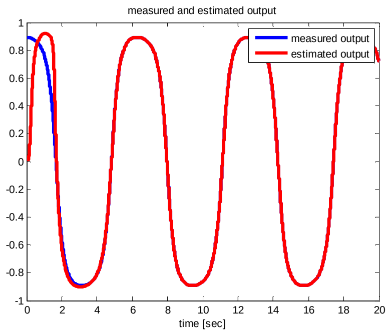

Fig. 1 The measured output $y=x+v$ and the estimated output $\hat{y}$ versus time

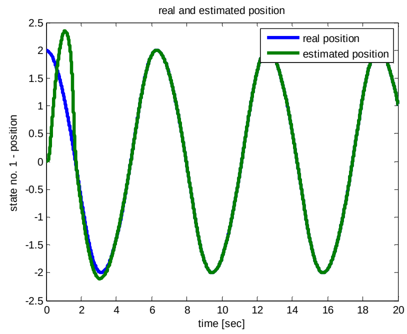

Fig. 2 The real position x and the estimated position state $\hat{x}$ versus time

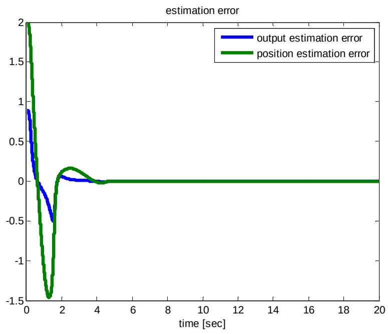

Fig. 3 The estimation errors versus time

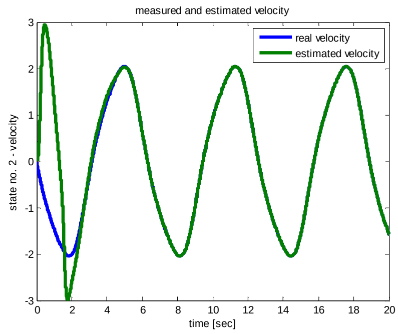

Fig. 4 The real velocity – $\dot{x}$ and the estimated velocity – $\hat{\dot{x}}$ versus time

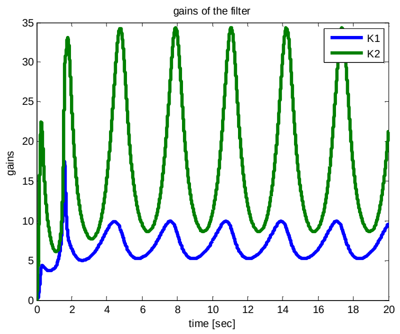

Fig. 5 The filter's gains K1- gain of the position state, K2 - gain of the velocity state versus time

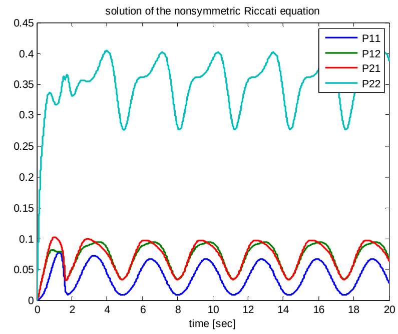

Fig. 6 The terms of the solution of the matrix Riccati equation - P versus time

With measurement noise and with system driving noise

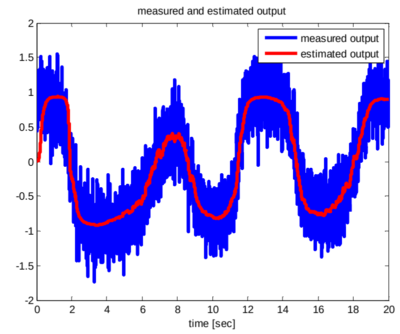

Fig. 7 The measured output, y=x+v and the estimated output $\hat{x}$ versus time

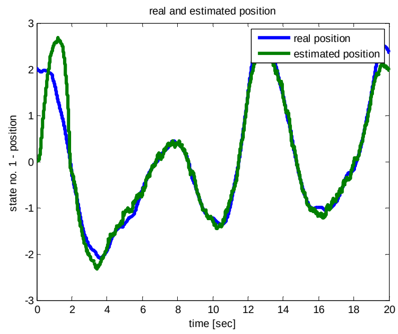

Fig. 8 The real position x and the estimated position state $\hat{x}$ versus time

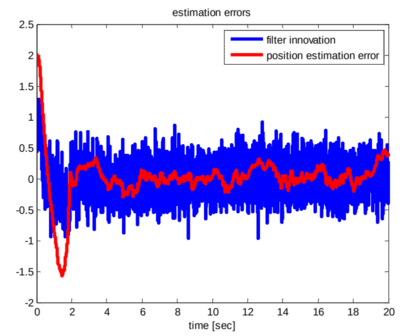

Fig. 9 The estimation errors versus time

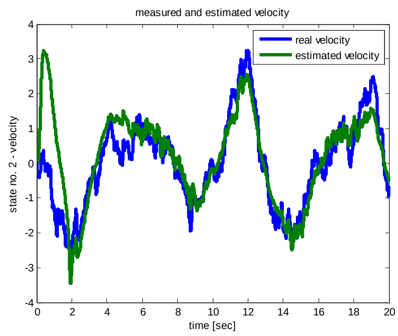

Fig. 10 The real velocity – $\dot{x}$ and the estimated velocity – $\hat{\dot{x}}$ versus time

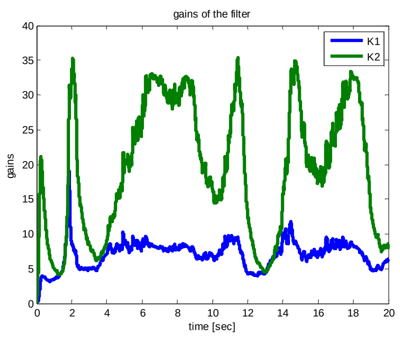

Fig. 11 The filter's gains, K1 gain of the position state, K2 gain of the velocity state versus time

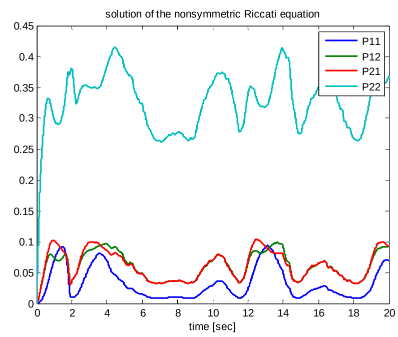

Fig. 12 The terms of the solution of the matrix Riccati equation - *P* versus time

CONCLUSION

The Joint Maximum Likelihood criterion was used to derive explicit solution of the optimal estimator for discrete nonlinear dynamic systems. Examples demonstrate it performance.

REFERENCES

[1] Särkkä S and Sarmavuori J, Gaussian Filtering and Smoothing for Continuous-Discrete Dynamic Systems, *Signal Processing*, September **2012.**
[2] Beneš VE, Exact Finite-Dimensional Filters for certain Diffusions with Nonlinear Drift, *Stochastics*, **1981**, Vol 5, p. 65-92.
[3] Daum F, Exact Finite Dimensional Nonlinear Filters, *IEEE Trans Autom Contr*, **1986,** Vol 31, No 7, p. 616-622
[4] Bryson AE and Frasier M, *Smoothing for Linear and Nonlinear Dynamical Systems*, Proc Optimum Systems Synthesis Conf Wright-Patterson Air Force Base, Ohio, February **1963.**
[5] Bryson AE and Ho Y, *Applied Optimal Control*, New York Hemisphere Publishing, **1975.**
[6] Jazwinski A H, *Stochastic Processes and Filtering Theory*, Academic Press, **1970.**
[7] Chan Z, Bayesian Filtering, From Kalman Filters to Particle Filters and Beyond, *Adaptive Syst Lab McMaster Univ*, Hamilton, ON, Canada .
[8] Xin M and Balakrishnan SN, A New Filtering Technique for a Class of Nonlinear Systems, Proceedings *of the 41st Conference on Decision and Control*, Las-Vegas, Nevada, USA, **2002.**
[9] Kalman RE, A New Ap.roach to Linear Filtering and Prediction Problems, *Trans ASME, Series D, Journal of Basic Engineering*, **1960,**Vol 82, p. 35-45.
[10] Kalman RE and Bucy RE, New Results to Linear Filtering and Prediction Theory, *Trans ASME, Journal of Basic Engineering*, **1961**, p. 95-108.
[11] Haessig DA and Friedland B, State Dependent Differential Riccati Equation for Nonlinear Estimation and Control, *2002 IFAC, 15th Triennial World Congress*, Barcelona, Spain, **2002.**
[12] Gelb A Ed, *Applied Optimal Estimation*, The MIT press, **1974.**
[13] Çimen T and Merttopçuoğlu A O, Asymptotically Optimal Nonlinear Filtering, Theory and Examples with Application to Target State Estimation, *Proceedings of the 17th World Congress The International Federation of Automatic Contro*l, Seoul, Korea, **2008**, p. 8611-8617.
[14] Denham WF and Pines S, Sequential Estimation when Measurement Nonlinearity is Comparable to Measurement Error, *AIAA J*ournal, **1966,** 4, 1071-1076.
[15] Morrison Norman, *Introduction to Sequential Smoothing and Prediction*, McGraw Hill **1969.**
[16] Rusnak I, Maximum Likelihood Optimal Estimator of Continuous Nonlinear Dynamic Systems, *IEEEI 2014, 28th Convention of Electrical and Electronics Engineers, Israel*, **2014**.
[17] Kwakernaak H and Sivan R, *Linear Optimal Control Systems*, Wiley-Interscience, **1972.**