Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Cramer-Rao Lower Bound for a Nonlinear Filtering Problem with Multiplicative Measurement Errors and Forcing Noise

Stepanov O.A.*, Vasilyev V.A.**

** Concern CSRI Elektropribor, JSC,*
National Research University of Information Technology, Mechanics and Optics,
49 Kronverkskiy pr., St. Petersburg, 197101, Russia, (e-mail: soalax@mail.ru)
*** Concern CSRI Elektropribor, JSC*
30, Malaya Posadskaya str., St. Petersburg, 197046, Russia, (e-mail: dolnakova1@yandex.ru)

Abstract: The recurrence algorithms for the Cramer-Rao lower bound for a discrete-time nonlinear filtering problem in the conditions when a forcing noise, measurement errors and initial covariance matrix depend on the state vector to be estimated are derived. It is assumed that the state vector being estimated includes a subvector of time-invariant unknown parameters. Some examples are given to illustrate the applicability of the algorithms obtained.

Keywords: nonlinear filtering, multiplicative measurement errors, multiplicative forcing noise unknown parameters, accuracy, Cramer lower bound.

1. INTRODUCTION

When efficient algorithms for processing of measurement data are developed in the context of the Bayesian filtering theory, it is a common practice for researchers to solve two problems: the problem of the analysis of the potential accuracy obtained using the algorithm, optimal in the sense of the chosen criterion, and the problem of design of a computationally economical algorithm that provides accuracy close to potential. Such an approach is widely used, in particular, for the development of algorithms for navigation data processing and tracking problems (Dmitriev and Stepanov 1998, Bergman 2001, Ristic et. al. 2004). The covariance matrix of estimation errors of the optimal algorithm is conventionally used as a characteristic of the potential accuracy. This matrix is determined by simulation which involves the procedure for calculating the optimal estimate. It is well known that, generally, it is impossible to design a universal and computationally convenient optimal algorithm for the problems of nonlinear filtering. Despite the fact that researchers have advanced in designing such algorithms recently due to, in particular, the application of various modifications of the Monte Carlo method (Doucet 2001, Gustafsson et al. 2002, Ristic 2004) the calculation of optimal estimates by these methods is computationally intensive (Snyder 2008, Stepanov and Berkovskiy 2011). In this regard, the development of approximate procedures for the analysis of potential accuracy of estimation is vitally important for the solution of applied problems. One of such procedures is based on the calculation of the Cramer-Rao lower bound (CRLB) (Van Trees 1968).

The methods of obtaining algorithms for CRLB calculation and their application in nonlinear filtering problems have been the subject matter of many publications (Galdos 1980, Van Trees and Bell 2007). For example, in (Koshaev and Stepanov 1997, Tichavsky et al. 1998, Simandl et al. 2001), the authors obtained convenient recurrence algorithms for CRLB calculation for discrete-time nonlinear filtering problems with additive measurement errors and forcing (process) noise in the equations for the state vector. These algorithms have been successfully used to solve a wide range of problems related in particular to the processing of navigation data (Dmitriev and Stepanov 1998, Bergman 1999, 2001, Batista et al 2013). However, in practice, there is often a need to solve problems in which the properties of forcing noise and measurement errors depend on the unknown state vector to be estimated, thus endowing them multiplicative nature. It is to this problem that the paper is devoted. Actually, we continue the research reported in (Stepanov et al. 2013). Here we suppose that not only properties of a forcing noise depend on the unknown state vector, but such dependency is also valid for measurement noise and the initial covariance matrix. These generalizations are very important in estimating the parameters of Markov random processes, widely used in the problems of navigation and tracking data filtering.

2. PROBLEM STATEMENT

Let us assume that we have composite $n+r$-dimensional vector $\tilde{x}_i = \left(x_i^T, \theta^T\right)^T$, which includes n-dimensional Markov sequence $x_i = \left(x_{i1}, x_{i2} \dots x_{in}\right)^T$ and r-dimensional vector $\theta = (\theta_1, \dots \theta_r)^T$ of unknown time-invariant parameters described by the following equations:

$$\left.\begin{aligned} x_i &= \Phi_i\left(x_{i-1},\theta\right) + \Gamma_i\left(x_{i-1},\theta\right) w_i, \\ \dot{\theta} &= \theta, \end{aligned}\right\} \tag{1}$$

and we also have m-dimensional measurements

$$y_i = s_i\left(x_i,\theta\right) + \Psi_i(\theta) v_i . \tag{2}$$

Here $\Phi_i(x_{i-1},\theta)$, $s_i(x_i,\theta)$ are the known nonlinear vector-functions of n and m dimensions; $\Gamma_i(x_{i-1},\theta)$, $\Psi_i(\theta)$ are the known matrices of $n\times p$ and $m\times m$ dimensions, the elements of which are nonlinear functions of their arguments; w_i and v_i are white-noise zero-mean Gaussian sequences of p and m dimensions, for which the relations $E\{w_l w_k^T\}=\delta_{lk}Q_l$, $E\{v_l v_k^T\}=\delta_{lk}R_l$ hold; Q_l and R_l are covariance matrices; δ_{lk} is the Kronecker operator; θ, x_0 are random vectors with the known probability density function (PDF) $f(x_0,\theta)=f(x_0/\theta)f(\theta)$, where $f(x_0/\theta)$ is Gaussian, i.e., $f(x_0/\theta)=N(x_0;0,P_0(\theta))$, with $E_{x_0}\{x_0\}=0$, $E_{x_0}\{x_0x_0^T\}=P_0(\theta)$. Along with $\tilde{x}_i=\left(x_i^T,\theta^T\right)^T$, we introduce composite vectors $X_i=\left(x_0^T,x_1^T,...,x_i^T\right)^T$, $\tilde{X}_i=\left(X_i^T,\theta^T\right)^T$, $Y_i=(y_1^T,...y_i^T)^T$ of $(i+1)n$, $((i+1)n+r)$, and im dimensions. We can write the following the Cramer-Rao inequality, for the vector $\tilde{X}_i=\left(X_i^T,\theta^T\right)^T$ (Galdos 1980):

$$\mathbf{J}_i^{-1}\le\mathbf{G}_i=E_{\tilde{X}_i,Y_i}\left\{\left(\tilde{X}_i-\hat{\tilde{X}}_i(Y_i)\right)\left(\tilde{X}_i-\hat{\tilde{X}}_i(Y_i)\right)^T\right\}, \quad (3)$$

where

$$\mathbf{J}_i=E_{\tilde{X}_i,Y_i}\left\{\frac{d\ln f\left(\tilde{X}_i,Y_i\right)}{d\tilde{X}_i}\left(\frac{d\ln f(\tilde{X}_i,Y_i)}{d\tilde{X}_i}\right)^T\right\}, \quad (4)$$

and $f(\tilde{X}_i,Y_i)$ is the PDF for vectors $\tilde{X}_i$ and Y_i. Let us separate the lower $(n+r)\times(n+r)$ diagonal block in $\mathbf{J}_i^{-1}$

$$\begin{bmatrix}\Upsilon_i^{x_i} & \Upsilon_i^{x_i\theta}\\ \Upsilon_i^{x_i\theta} & \Upsilon_i^{\theta}\end{bmatrix}=\left(\mathbf{0}_{(n+r)\times(in)},\mathbf{I}_{(n+r)}\right)\mathbf{J}_i^{-1}\begin{pmatrix}\mathbf{0}_{(in)\times(n+r)}\\ \mathbf{I}_{(n+r)}\end{pmatrix}, (5)$$

where $\mathbf{0}_{(n+r)\times(in)}$ – $(n+r)\times(in)$ is a zero matrix and $\mathbf{I}_{(n+r)}$ is a unit $n+r$ matrix. The matrices $\Upsilon_i^{x_i}$, Υ_i^{θ} determine CRLB for vectors x_i and θ.

The purpose of this work is to obtain a recurrence algorithm for Υ_i^{θ}, $\Upsilon_i^{x_i}$.

3. ALGORITHM FOR CRLB

Doing mathematical operations in the way similar to that of (Stepanov et al. 2013), we can show that for Υ_i^{θ} and $\Upsilon_i^{x_i}$ the following relations hold good:

$$\Upsilon_i^{\theta}=\left(\tilde{F}_i^{\theta}-\tilde{L}_i\left(\tilde{\Xi}_i\right)^{-1}\tilde{L}_i^T\right)^{-1}, \quad (6)$$

$$\Upsilon_i^{x_i}=\tilde{\Xi}_i^{-1}+\tilde{\Xi}_i^{-1}\tilde{L}_i\Upsilon_i^{\theta}\tilde{L}_i^T\tilde{\Xi}_i^{-1}. \quad (7)$$

Here

$$\left.\begin{aligned}&\tilde{F}_i^{\theta}=\tilde{F}_{i-1}^{\theta}+\delta s_i^{\theta}+\delta\Phi_i^{\theta}+\delta\tilde{Q}_i^{\theta}+\delta\tilde{R}_i^{\theta}-\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}+\delta\tilde{Q}_i^{x_{i-1},\theta}\right)\times\\&\times\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}+\delta\tilde{Q}_i^{x_{i-1}}\right)^{-1}\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}+\delta\tilde{Q}_i^{x_{i-1},\theta}\right)^T,\\&\tilde{F}_0^{\theta}=\tilde{F}_{01}^{\theta}+\tilde{F}_{02}^{\theta},\end{aligned}\right\}(8)$$

$$\tilde{L}_i=\delta s_i^{x_j,\theta}+\delta\tilde{R}_i^{x_i,\theta}-\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}+\delta\tilde{Q}_i^{x_{i-1},\theta}\right)\times$$
$$\times\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}+\delta\tilde{Q}_i^{x_{i-1}}\right)^{-1}\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}+\delta\tilde{Q}_i^{x_{i-1},\theta}\right)^T,\tilde{L}_0=0, \quad (9)$$

$$\tilde{\Xi}_i=\delta s_i^{x_i}+\delta\tilde{R}_i^{x_i}+\bar{\tilde{Q}}_i^{-1}-\bar{\Phi}_i^T\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}+\delta\tilde{Q}_i^{x_{i-1}}\right)^{-1}\bar{\Phi}_i, (10)$$
$$\tilde{\Xi}_0=E_{\theta}\left(P_0^{-1}(\theta)\right).$$

Here:

$$\tilde{F}_{01}^{\theta}=E_{\theta}\left(\frac{d}{d\theta}\ln f(\theta)\left(\frac{d}{d\theta}\ln f(\theta)\right)^T\right),$$

$$\tilde{F}_{02}^{\theta}=\frac{1}{2}E_{\theta}\left(tr\left(\frac{d\left(P_0(\theta)\right)^{-1}}{d\theta_l}P_0(\theta)\times\frac{d\left(P_0(\theta)\right)^{-1}}{d\theta_{\mu}}P_0(\theta)\right)^T\right)-$$
$$-\frac{1}{2}E_{\theta}\begin{pmatrix}\dfrac{d\ln f(\theta)}{d\theta}\left(\dfrac{d\ln\det\left(P_0(\theta)\right)}{d\theta}\right)^T+\\+\dfrac{d\ln\det\left(P_0(\theta)\right)}{d\theta}\left(\dfrac{d\ln f(\theta)}{d\theta}\right)^T\end{pmatrix},$$

$$\bar{\Phi}_i=E_{x_{i-1},\theta}\left(\frac{\partial\Phi_i(x_{i-1},\theta)}{\partial x_{i-1}^T}\tilde{Q}_i^{-1}(x_{i-1},\theta)\right),$$

$$\delta s_i^{x_i}=E_{x_i,\theta}\left(\frac{\partial s_i^T(x_i,\theta)}{\partial x_i}\tilde{R}_i^{-1}(\theta)\frac{\partial s_i(x_i,\theta)}{\partial x_i^T}\right),$$

$$\delta s_i^{\theta}=E_{x_i,\theta}\left(\frac{\partial s_i^T(x_i,\theta)}{\partial\theta}\tilde{R}_i^{-1}(\theta)\frac{\partial s_i(x_i,\theta)}{\partial\theta^T}\right),$$

$$\delta s_i^{x_i,\theta}=E_{x_i,\theta}\left(\frac{\partial s_i^T(x_i,\theta)}{\partial\theta}\tilde{R}_i^{-1}(\theta)\frac{\partial s_j(x_i,\theta)}{\partial x_i^T}\right),$$

$$\delta\Phi_i^{x_{i-1}}=E_{x_{i-1},\theta}\left(\frac{\partial\Phi_i^T(x_{i-1},\theta)}{\partial x_{i-1}}\tilde{Q}_i^{-1}(x_{i-1},\theta)\frac{\partial\Phi_i(x_{i-1},\theta)}{\partial x_{i-1}^T}\right),$$

$$\delta\Phi_i^{\theta}=E_{x_{i-1},\theta}\left(\frac{\partial\Phi_i^T(x_{i-1},\theta)}{\partial\theta}\tilde{Q}_i^{-1}(x_{i-1},\theta)\frac{\partial\Phi_i(x_{i-1},\theta)}{\partial\theta^T}\right),$$

$$\delta\Phi_i^{x_{i-1},\theta}=E_{x_{i-1},\theta}\left(\frac{\partial\Phi_i^T(x_{i-1},\theta)}{\partial\theta}\tilde{Q}_i^{-1}(x_{i-1},\theta)\frac{\partial\Phi_i(x_{i-1},\theta)}{x_{i-1}^T}\right),$$

$$\tilde{Q}_i(x_{i-1},\theta)=\Gamma_i(x_{i-1},\theta)Q_i\Gamma_i^T(x_{i-1},\theta),$$

$$\bar{\tilde{Q}}_i^{-1}=E_{x_{i-1},\theta}\left(\tilde{Q}_i^{-1}(x_{i-1},\theta)\right),\ \tilde{R}_i(\theta)=\Psi_i(\theta)R_i\Psi_i^{\mathrm{T}}(\theta),$$

$$\delta\tilde{R}_i^{\theta}(\theta)=\frac{1}{2}E_{x_{i-1},\theta}\left\{tr\left(\frac{\partial\tilde{R}_i^{-1}(\theta)}{\partial\theta_l}\tilde{R}_i(\theta)\frac{\partial\tilde{R}_i^{-1}(\theta)}{\partial\theta_{\mu}}\tilde{R}_i(\theta)\right)\right\},$$

$$l,\mu=\overline{1.r},$$

$$\delta\tilde{R}_i^{x_i}=\left\{\delta\tilde{R}_i^{x_i}(l,\mu)\right\}=\frac{1}{2}E_{x_i,\theta}\left\{tr\left(\frac{\partial\tilde{R}_i^{-1}(x_i,\theta)}{\partial x_{i-1,l}}\tilde{R}_i(x_i,\theta)\frac{\partial\tilde{R}_i^{-1}(x_i,\theta)}{\partial x_{i-1,\mu}}\tilde{R}_i(x_i,\theta)\right)\right\}\quad l,\mu=\overline{1.n},$$

$$\delta\tilde{R}_i^{x_i,\theta}=\left\{\delta\tilde{R}_i^{x_i,\theta}(l,\mu)\right\}=\frac{1}{2}E_{x_i,\theta}\left\{tr\left(\frac{\partial\tilde{R}_i^{-1}(x_i,\theta)}{\partial\theta_l}\tilde{R}_i(x_i,\theta)\frac{\partial\tilde{R}_i^{-1}(x_i,\theta)}{\partial x_{i,\mu}}\tilde{R}_i(x_i,\theta)\right)\right\}\quad \mu=\overline{1.n}, l=\overline{1.r},$$

$$\delta\tilde{Q}_i^{x_{i-1}}\{l,\mu\}=\frac{1}{2}E_{x_{i-1},\theta}\left\{tr\begin{pmatrix}\dfrac{\partial\tilde{Q}_i^{-1}(x_{i-1},\theta)}{\partial x_{i-1,l}}\tilde{Q}_i(x_{i-1},\theta)\times\\ \times\dfrac{\partial\tilde{Q}_i^{-1}(x_{i-1},\theta)}{\partial x_{i-1,\mu}}\tilde{Q}_i(x_{i-1},\theta)\end{pmatrix}\right\}, l,\mu=\overline{1.n},$$

$$\delta\tilde{Q}_i^{\theta}\{l,\mu\}=\frac{1}{2}E_{x_{i-1},\theta}\left\{tr\begin{pmatrix}\dfrac{\partial\tilde{Q}_i^{-1}(x_{i-1},\theta)}{\partial\theta_l}\tilde{Q}_i(x_{i-1},\theta)\times\\ \times\dfrac{\partial\tilde{Q}_i^{-1}(x_{ji-1},\theta)}{\partial\theta_\mu}\tilde{Q}_i(x_{i-1},\theta)\end{pmatrix}\right\}, l,\mu=\overline{1.r},$$

$$\delta\tilde{Q}_i^{x_{i-1},\theta}\{l,\mu\}=\frac{1}{2}E_{x_{i-1},\theta}\left\{tr\begin{pmatrix}\dfrac{\partial\tilde{Q}_i^{-1}(x_{i-1},\theta)}{\partial\theta_l}\tilde{Q}_i(x_{i-1},\theta)\times\\ \times\dfrac{\partial\tilde{Q}_i^{-1}(x_{i-1},\theta)}{\partial x_{i-1,\mu}}\tilde{Q}_i(x_{i-1},\theta)\end{pmatrix}\right\}, \mu=\overline{1.n}, l=\overline{1.r}.$$

In these equations the following notations are used:

$$\frac{ds(x)}{dx^T}=\begin{bmatrix}\dfrac{\partial s_1(x)}{\partial x_1} & . & . & . & \dfrac{\partial s_1(x)}{\partial x_n}\\ . & . & . & . & .\\ . & . & . & . & .\\ . & . & . & . & .\\ \dfrac{\partial s_m(x)}{\partial x_1} & . & . & . & \dfrac{\partial s_m(x)}{\partial x_n}\end{bmatrix},\ \frac{ds^T(x)}{dx}=\left[\frac{ds(x)}{dx^T}\right]^T.$$

Note that when $f(\theta)=N(\theta;0,P_\theta)$ and $P_0(\theta)=P_0$ we can write $\tilde{F}_0^\theta=P_\theta^{-1}$. Let us consider some specific cases.

Case 1. The subvector θ is absent ($r=0$), the measurement errors and forcing noise are additive; moreover, $\Gamma_i(x_{i-1})=\Gamma_i$, $\Psi_i(\theta)=I$ is a unit matrix, $f(x_0/\theta)=f(\theta)$, $P_0(\theta)=P_0$, $f(\theta)=N(\theta;0,P_\theta)$. For these assumptions $\delta\tilde{R}_i^\theta=0$, $\delta\tilde{Q}_i^\theta=0$, $\delta\tilde{Q}_i^{x_{i-1},\theta}=0$, $\delta\tilde{Q}_i^{x_{i-1}}=0$, $\bar{\tilde{Q}}_i=\tilde{Q}_i=\Gamma_iQ_i\Gamma_i^T$, $\delta s_i^\theta=0$, $\delta\Phi_i^\theta=0$, $\delta s_i^{x_j,\theta}=0$, $\delta\Phi_i^{x_{i-1},\theta}=0$, $\tilde{L}_j=0$, $j=\overline{0,i-1}$ and therefore

$$\Upsilon_i^{x_i}=\tilde{\Xi}_i^{-1}, \tag{11}$$

$$\tilde{\Xi}_i=\delta s_i^{x_i}+\tilde{Q}_i^{-1}-\bar{\Phi}_i^T\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}\right)^{-1}\bar{\Phi}_i,\ \tilde{\Xi}_0=P_0^{-1}. \tag{12}$$

Since $\tilde{Q}_i$ does not depend on the state vector, we can write:

$$\bar{\Phi}_i=E_{x_{i-1},\theta}\left(\frac{\partial\Phi_i(x_{i-1},\theta)}{\partial x_{i-1}^T}\right)\tilde{Q}_i^{-1}=\bar{\Phi}_i^*\tilde{Q}_i^{-1},\text{ and}$$

$$\tilde{\Xi}_i=\delta s_i^{x_i}+\tilde{Q}_i^{-1}-\tilde{Q}_i^{-1}(\bar{\Phi}_i^*)^{\mathrm{T}}\left(\tilde{\Xi}_{i-1}+\delta\Phi_i\right)^{-1}\bar{\Phi}_i^*\tilde{Q}_i^{-1},\tilde{\Xi}_0=P_0^{-1}, \tag{13}$$

where $\bar{\Phi}_i^*=E_{x_{i-1}}\left(\dfrac{d\Phi_i(x_{i-1})}{dx_i^T}\right)$,

$\delta\Phi_i^{x_{i-1}}=E_{x_{i-1}}\left\{\dfrac{\Phi_i^T(x_{i-1})}{dx_{i-1}}\tilde{Q}_i^{-1}\dfrac{\Phi_i(x_{i-1})}{dx_{i-1}^T}\right\}$, $\delta s_i^{x_i}=E_{x_i}\left\{\dfrac{s_i^T(x_i)}{dx_i}R_i^{-1}\dfrac{s_i(x_i)}{dx_i^T}\right\}$.

It is clear that (13) coincide with the equations in (Koshaev et al. 1997, Tichavsky 1998, Simandl et al. 2001).

Let us rewrite (12) in the form:

$$\left.\begin{aligned}\tilde{\Xi}_i&=\delta s_i^{x_i}+\tilde{Q}_i^{-1}-\tilde{Q}_i^{-1}(\bar{\Phi}_i^*)^{\mathrm{T}}\times\\ &\times\left((\bar{\Phi}_i^*)^{\mathrm{T}}\tilde{Q}_i^{-1}\bar{\Phi}_i^*+\tilde{\Xi}_{i-1}+\Delta\Phi_i\right)^{-1}\bar{\Phi}_i^*\tilde{Q}_i^{-1}\end{aligned}\right\}, \tag{14}$$

where $\Delta\Phi_i=\delta\Phi_i^{x_{i-1}}-\bar{\Phi}_i^*\tilde{Q}_i^{-1}(\bar{\Phi}_i^*)^{\mathrm{T}}$.

Applying the matrix inversion lemma to (14), we can write one more variant of the recurrence relation for $\tilde{\Xi}_i$

$$\tilde{\Xi}_i=\delta s_i^{x_i}+\left((\bar{\Phi}_i^*)^{\mathrm{T}}(\tilde{\Xi}_{i-1}+\Delta\Phi_i)^{-1}\bar{\Phi}_i^*+\tilde{Q}_i\right)^{-1}. \tag{15}$$

Remark. It should be noted that by including the subvector θ in the state vector x_i, we can also obtain the CRLB for θ using (15) for some cases (Koshaev and Steanov 1997), which, for example, is true when the equation for the state-vector is linear. It is clear because $\Delta\Phi_i=0$ and using (15), we do not need to invert matrix $\tilde{Q}_i$. However we should keep it in mind that in a general case the CRLB for θ obtained using the algorithm with the vector of unknown parameters included in the state vector is higher or equal to the CRLB for θ calculated using the algorithm obtained in this paper (Koshaev 1998). In other words, the CRLB for θ obtained using the algorithm considered in this paper is more exact. For more details, see variant 2, for example.

Case 2. The state vector includes subvector θ ($r>0$) and all additional assumptions are the same as for case 1, then $\tilde{F}_0^\theta=P_\theta^{-1}$, $\delta\tilde{R}_i^\theta=0$, $\delta\tilde{Q}_i^\theta=0$, $\delta\tilde{Q}_i^{x_{i-1},\theta}=0$, $\delta\tilde{Q}_i^{x_{i-1}}=0$, $\bar{\tilde{Q}}_i^{-1}=\tilde{Q}_i^{-1}$. Equations (6), (7) are the same, but the equations for the matrices included in them are different, i.e.,

$$\Upsilon_i^\theta=\left(\tilde{F}_i^\theta-\tilde{L}_i\left(\tilde{\Xi}_i\right)^{-1}\tilde{L}_i^T\right)^{-1},\qquad \Upsilon_i^{x_i}=\tilde{\Xi}_i^{-1}+\tilde{\Xi}_i^{-1}\tilde{L}_i\Upsilon_i^\theta\tilde{L}_i^T\tilde{\Xi}_i^{-1},$$

$$\tilde{F}_i^\theta=\tilde{F}_{i-1}^\theta+\delta s_i^\theta+\delta\Phi_i^\theta-\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}\right)\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}\right)^{-1}\times$$

$$\times\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}\right)^T,\ \tilde{F}_0^\theta=E_\theta\left(\frac{d}{d\theta}\ln f(\theta)\left(\frac{d}{d\theta}\ln f(\theta)\right)^T\right),$$

$$\tilde{L}_i=\delta s_i^{x_j,\theta}-\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}\right)\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}\right)^{-1}\times$$

$$\times\left(\tilde{L}_{i-1}+\delta\Phi_i^{x_{i-1},\theta}\right)^T,\ \tilde{L}_0=0,$$

$$\tilde{\Xi}_i=\delta s_i^{x_i}+\tilde{Q}_i^{-1}-\bar{\Phi}_i^T\left(\tilde{\Xi}_{i-1}+\delta\Phi_i^{x_{i-1}}\right)^{-1}\bar{\Phi}_i,\ \tilde{\Xi}_0=P_0^{-1}.$$

If, in addition, $\delta s_i^{x_j,\theta} = 0$ and $\delta\Phi_i^{x_{i-1},\theta} = 0$, then all $\tilde{L}_j = 0$, $j = \overline{0,i-1}$; therefore,

$$\Upsilon_i^\theta = \left(\tilde{F}_i^\theta\right)^{-1}, \Upsilon_i^{x_i} = \tilde{\Xi}_i^{-1}, \tag{16}$$

$$\tilde{F}_i^\theta = \tilde{F}_{i-1}^\theta + \delta s_i^\theta + \delta\Phi_i^\theta, \tag{17}$$

$$\tilde{\Xi}_i = \delta s_i^{x_i} + \tilde{Q}_i^{-1} - \bar{\Phi}_i^T \left(\tilde{\Xi}_{i-1} + \delta\Phi_i^{x_{i-1}}\right)^{-1} \bar{\Phi}_i \text{ , } \tilde{\Xi}_0 = P_0^{-1} \text{ ,} \tag{18}$$

or

$$\tilde{\Xi}_i = \delta s_i + \left((\bar{\Phi}_i^*)^{\mathrm{T}} (\tilde{\Xi}_{i-1} + \Delta\Phi_i)^{-1} \bar{\Phi}_i^* + \tilde{Q}_i\right)^{-1}, \tilde{\Xi}_0 = P_0^{-1}. \tag{19}$$

Below, we give some simplest examples to illustrate the application of the relations obtained.

4. EXAMPLE

Assume that we need to estimate an unknown parameter $\theta = q$ of a random walk (Wiener process) $z(t)$ by its discrete measurements with additive measurement errors. Let us consider different variants of the problem solution.

Model 1. We use the following model for discrete time:

$$\left.\begin{aligned} x_i &= x_{i-1} + \sqrt{\Delta t}\, w_i, \\ \dot{q} &= 0, \end{aligned}\right\} \tag{20}$$

$$y_i = z_i + v_i = s_i(x_i, q) + v_i = q x_i + v_i . \tag{21}$$

where $\Phi_i\left(x_{i\text{-}1}, q\right) = x_{i\text{-}1}$, $s_i(x_i, q) = q x_i$, $\Gamma_i = \sqrt{\Delta t}\sigma_w$, Δt is the sampling interval; v_i and w_i are zero-mean Gaussian white noise with variances σ_v^2 and σ_w^2, respectively; x_0 and q are independent random values with PDF $f(x_0) = N(x_0; 0, \sigma^2)$ and $f(q)$ is a PDF, for which the $\bar{q} = E(q)$ and $\sigma_q^2 = E(q - \bar{q})^2$ are known. A feature of this model is that the shaping filter for z_i does not depend on θ and nonlinearity is only due to nonlinearity in measurements. Using the above relations, we can write:

$$\tilde{Q}_i^{-1}(x_{i\text{-}1}, q) = \frac{1}{\sigma_w^2 \Delta t}, \bar{\Phi}_j = \frac{1}{\sigma_w^2 \Delta t}, \delta s_i^{x_i} = E_q\left(\frac{q^2}{\sigma_v^2}\right) = \frac{\bar{q}^2 + \sigma_q^2}{\sigma_v^2},$$

$$\delta s_i^q = E_{x_i}\left(\frac{x_i^2}{\sigma_v^2}\right) = \frac{i\Delta t \sigma_w^2 + \sigma^2}{\sigma_v^2}, \quad \delta s_i^{x_i,q} = 0, \quad \delta\Phi_i^{x_{i-1}} = \frac{1}{\sigma_w^2 \Delta t},$$

$\delta\Phi_i^q = 0$, $\delta\Phi_i^{x_{i-1},q} = 0$, $\delta\tilde{R}_i^\theta = 0$, $\delta\tilde{Q}_i^{x_{i-1}} = 0$, $\delta\tilde{Q}_i^q = 0$, $\delta\tilde{Q}_i^{x_{i-1},q} = 0$, $\bar{\tilde{Q}}_i^{-1} = (\sigma_w^2 \Delta t)^{-1}$, $\bar{\Phi}_i^* = 1$, $\Delta\Phi_i = 0$, $L_j = 0, j = \overline{0.i}$.

Taking into consideration the fact that this example corresponds to case 2, and, in addition, all $L_j = 0$, $j = \overline{0.i}$, we can use (16), (17), and (19). Thus:

$$\tilde{F}_i^q = \tilde{F}_{i-1}^q + \delta s_i^q = \tilde{F}_{i-1}^q + \frac{i\Delta t \sigma_w^2 + \sigma^2}{\sigma_v^2}, \quad \tilde{F}_0^q = \tilde{F}_{01}^q,$$

$$\tilde{\Xi}_i = \delta s_i^{x_i} + \left((\tilde{\Xi}_{i-1})^{-1} + \tilde{Q}_i\right)^{-1} =$$

$$\tilde{\Xi}_i = \frac{\bar{q}^2 + \sigma_q^2}{\sigma_v^2} + \left((\tilde{\Xi}_{i-1})^{-1} + \sigma_w^2 \Delta t\right)^{-1}, \tilde{\Xi}_0 = 1/\sigma^2.$$

Finally,

$$\Upsilon_i^q = \left(\tilde{F}_0^q + \frac{0.5(i+1) i \Delta t \sigma_w^2 + i\sigma^2}{\sigma_v^2}\right)^{-1}, \tag{22}$$

$$\Upsilon_i^{x_i} = \left(\left(\Upsilon_{i-1}^{x_{i-1}} + \Delta t \sigma_w^2\right)^{-1} + \frac{\sigma_q^2 + \bar{q}^2}{\sigma_v^2}\right)^{-1}, \ \Upsilon_0^{x_0} = \sigma^2 . \tag{23}$$

It is also easy to see that for $f(q) = N(q; \bar{q}, \sigma_q^2)$, then $\tilde{F}_0^q = \frac{1}{\sigma_q^2}$. Thus, we can state the fact that in the case under consideration, the type of $f(q)$ at fixed values of $\bar{q} = E(q)$ and $\sigma_q^2 = E(q - \bar{q})^2$ does not practically affect the final result.

It should be noted that the CRLB for the model (20), (21) is equivalent to the covariance in the linear estimation problem of vector (20) by measurements

$$y_{i1}^{lin} = \left(\sqrt{\sigma_q^2 + \bar{q}^2}\right) x_i + v_{i1}^{lin},$$

$$y_{i2}^{lin} = \left(\sqrt{\sigma^2 + i\Delta t \sigma_w^2}\right) q + v_{i2}^{lin},$$

where v_{i1}^{lin} are independent of x_0 and q zero-mean Gaussian white-noise sequences with variances σ_v^2, whereas x_0 and q are independent of each other. Gaussian random values with variances σ^2 and $\left(\tilde{F}_0^q\right)^{-1}$. In other words, the value that determines the CRLB for q corresponds to the case of q estimation from measurements of the form (21) under the assumption that x_i is replaced by the known coefficient $\left(\sqrt{\sigma^2 + i\Delta t \sigma_w^2}\right)$. In turn, the CRLB for x_i corresponds to the case of x_i estimation from the same measurements, but under another assumption, namely, that q is replaced by the known coefficient $\left(\sqrt{\sigma_q^2 + \bar{q}^2}\right)$.

Model 2. Let us include the unknown parameters θ in the state vector $x_i^* = (x_i, q)^T$ and use the same model (20), (21), but in so doing, our aim is to find the recurrence relation for the CRLB for vector x_i^*. In this case, taking into consideration the fact that $s_i(x_i^*) = s_i(x_i, q) = q x_i$ $\tilde{\Phi}_i\left(x_i^*\right) = \begin{bmatrix} \Phi_i\left(x_{i\text{-}1}, q\right) \\ q \end{bmatrix} = \begin{bmatrix} x_{i-1} \\ q \end{bmatrix}$ and $\tilde{\Gamma}_i = \begin{bmatrix} \Gamma_i \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{\Delta t}\sigma_w \\ 0 \end{bmatrix}$, we can use (15) which does not require nonsingularity of matrix $\tilde{Q}_i$. Since $\dfrac{ds_i(x_i^*)}{d(x_i^*)^T} = \left(\dfrac{\partial s_i(x_i^*)}{\partial x_{i1}^*}, \dfrac{\partial s_i(x_i^*)}{\partial x_{i2}^*}\right) = (q, x_i)$, then

$$\delta s_i^{x_i^*} = E_{\tilde{x}_i} \frac{1}{\sigma_v^2}\left\{\begin{bmatrix} q^2 & x_i q \\ x_i q & x_i^2 \end{bmatrix}\right\} =$$

$$= \frac{1}{\sigma_v^2}\left\{\begin{bmatrix} \sigma_q^2 + \bar{q}^2 & 0 \\ 0 & \sigma^2 + i\Delta t \sigma_w^2 \end{bmatrix}\right\}.$$

By virtue of the fact that $\Delta\tilde{\Phi}_i = 0$ and $\bar{\tilde{\Phi}}_i^* = E$, (19) takes the form:

$$\Upsilon_i^{x_i^*} = \left(\delta s_i^{x_i^*} + \left(\Upsilon_{i-1}^{x_i^*} + \tilde{\Gamma}_i \tilde{\Gamma}_i^T\right)^{-1}\right)^{-1} =$$

$$= \left(\frac{1}{\sigma_v^2}\left\{\begin{bmatrix} \sigma_q^2 + \bar{q}^2 & 0 \\ 0 & \sigma^2 + i\Delta t \sigma_w^2 \end{bmatrix}\right\} + \left(\Upsilon_{i-1}^{x_i^*} + \begin{bmatrix} \Delta t \sigma_w^2 & 0 \\ 0 & 0 \end{bmatrix}\right)^{-1}\right)^{-1}, \tag{24}$$

where

$$\Upsilon_0^{x_0^*} = P_0 = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \left(\tilde{F}_0^q\right)^{-1} \end{bmatrix}.$$

It is easy to see that the result generating by (24) coincide with (22) and (23). Taking into consideration the above remark, we note that in this example, the CRLBs corresponding to different algorithms are identical.

Model 3. We can use another shaping filter for $z_i = x_i$:

$$x_i = x_{i-1} + q\sqrt{\Delta t}\, w_i, \qquad \dot{q} = 0, \tag{25}$$

$$y_i = x_i + v_i, \tag{26}$$

where v_i and w_i x_0 and q are the same as in the previous case. As in the first two models, we assume that for $f(q)$, the first two moments $\bar{q} = E(q)$ and $\sigma_q^2 = E(q-\bar{q})^2$ are known, and, besides, the value of $a = E_q\left(\dfrac{1}{q^2}\right) = \int \dfrac{1}{q^2} f(q)dq$ is also determined. It should be noted that Gaussian PDF does not satisfy the latter requirement because such integral diverges. The feature of this statement is that the model for the measurements are linear, whereas equation for the state vector is nonlinear, since the coefficient of the forcing noise depends on the unknown parameter. From (25)–(26) it follows that $\Phi_i(x_{i-1},\theta) = x_{i-1}$, $s_i(x_i, q) = x_i$, $\Gamma_i(x_{i-1}, q) = q$. In this case:

$\tilde{Q}_i^{-1}(x_{i-1}, q) = \dfrac{1}{q^2\sigma_w^2\Delta t}$, $\bar{\Phi}_i = E_q\left(\dfrac{1}{q^2\sigma_w^2\Delta t}\right) = a$, $\delta s_i^{x_i} = \dfrac{1}{\sigma_v^2}$, $\delta s_i^q = 0$, $\delta s_i^{x_i,q} = 0$, $\delta\Phi_i^{x_{i-1}} = E_q\left(\dfrac{1}{q^2\sigma_w^2\Delta t}\right) = a$, $\delta\Phi_i^q = 0$, $\delta\Phi_i^{x_{i-1},q} = 0$, $\tilde{Q}_i(x_{i-1},q) = \Gamma_i(x_{i-1},q) Q_i \Gamma_i^T(x_{i-1},q) = q^2\Delta t\sigma_w^2$, $\delta\tilde{R}_i^\theta = 0$, $\delta\tilde{Q}_i^{x_{i-1}} = 0$, $\delta\tilde{Q}_i^q = \dfrac{1}{2}E_q\left(\dfrac{1}{q^2\sigma_w^2\Delta t}\right) = \dfrac{a}{2}$, $\delta\tilde{Q}_i^{x_{i-1},q} = 0$, $\bar{\tilde{Q}}_i^{-1} = E_q\left(\dfrac{1}{q^2\sigma_w^2\Delta t}\right) = a$, $\bar{\Phi}_i^* = 1$, $L_j = 0$, $j = \overline{0.i}$, $\tilde{F}_0^q = E_q\left(\dfrac{d}{dq}\ln f(q)\left(\dfrac{d}{dq}\ln f(q)\right)^T\right)$.

Since here, too, all $L_j = 0$, $j = \overline{0.i}$, using (19) we can write

$$\tilde{F}_i^q = \tilde{F}_{i-1}^q + \frac{a}{2},\ \tilde{F}_0^q = E_q\left(\frac{d}{dq}\ln f(q)\left(\frac{d}{dq}\ln f(q)\right)^T\right),$$

$$\tilde{\Xi}_i = \frac{1}{\sigma_v^2} + a - \frac{a^2}{\tilde{\Xi}_{i-1} + a} = \frac{\left(\tilde{\Xi}_{i-1}(a\sigma_v^2 + 1) + a\right)}{\sigma_v^2(\tilde{\Xi}_{i-1} + a)}, \tilde{\Xi}_0 = P_0^{-1}.$$

Therefore, we have

$$\Upsilon_q = \left(\tilde{F}_0^q + \frac{ia}{2}\right)^{-1}, \tag{27}$$

$$\Upsilon_i^{x_i} = \frac{\sigma_v^2\left(\left(\Upsilon_{i-1}^{x_{i-1}}\right)^{-1} + a\right)}{\left(\Upsilon_{i-1}^{x_{i-1}}\right)^{-1}\left(a\sigma_v^2 + 1\right) + a},\ \Upsilon_0^{x_0} = P_0. \tag{28}$$

In this example we calculate the CRLB for q, which determines the properties of the Wiener process. It is interesting to compare (22) and (27) for the same values of $\bar{q}$ and σ_q^2. Let assume that $f(q) = \dfrac{3q^2 e^{-0.5q^3\sigma_q^{-3}}}{\sigma_q^3}$ (Weibull PDF). In this case, $\sigma_q = \dfrac{3\sqrt{3}\Gamma\left(\frac{2}{3}\right)}{2\pi}\bar{q}$, $a = \dfrac{2\pi}{\sqrt{3}\sigma_q^2\Gamma\left(\frac{2}{3}\right)}$, $\tilde{F}_0^q = \dfrac{4a}{3}$, where $\Gamma(\bullet)$ is the Gamma function. Fig. 1 presents the results of the CRLB calculations obtained using (22) and (27), with $\sigma_q = \sqrt{0.0019}$, $\bar{q} = 0.12$, $\sigma_v = 0.3$ $\sigma = 1.3$, $\sigma_w = 1$, and $\Delta t = 1$.

Here, we also give the values of the root-mean-square (RMS) error for the optimal estimate computed using Monte-Carlo simulation as $\sigma_q^{MC}(i) \approx \dfrac{1}{L}\sum_{j=1}^{L}\left(q^j - \hat{q}_i^j\right)^2$, where L-number of samples in Monte-Carlo simulation; q^j and $\hat{q}_i^j$ are the samples and optimal estimates calculated using the algorithm described, for example, in (Ivanov et al., 2000).

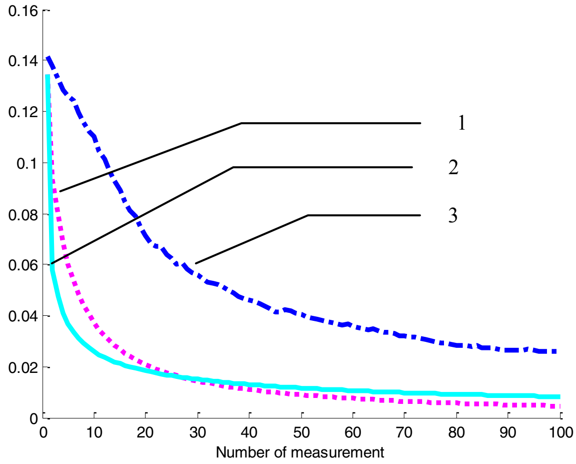

Fig. 1. 1 – CRLB for model (25), (26); 2 – CRLB for model (20), (21), 3 – RMS error for the optimal estimate for two models, L=500.

From the curves above it follows that, firstly, the CRLB is, unfortunately, significantly less than the real RMS error and, secondly, the CRLB depends on the model used for z_i. It is clear that in the accuracy analysis it makes sense to use an upper envelope corresponding to the two CRLBs.

5. CONCLUSIONS

Recurrence relations have been obtained for the calculation of CRLB in the discrete-time nonlinear filtering problem in the conditions when the forcing noise, measurement errors and initial covariance matrix depend on the state vector to be estimated which also includes the subvector of unknown time-invariant parameters.

Some specific cases have been considered. The relation between the derived recurrence algorithm for the CRLB calculation and the known algorithms corresponding to the case of additive forcing and measurement noise has been established.

An example of CRLB calculation in the estimation problem of the parameters of the random walk process has been considered. The results obtained allowed the conclusion that there is an obvious dependence of CRLB on the type of the model used to describe the process under study for nonlinear filtering problem. This dependence is worth further study.

REFERENCES

Batista, P., Silvestre C., Oliveira P. (2013). Preliminary results on the estimation performance of single range source localization. *Proceedings of 21st Mediterranean Conference on Control & Automation.* Platanias-Chania, Crete, Greece, June 25-28, pp. 421–424.

Bergman N. (1999). *Recursive Bayesian estimation. Navigation and tracking applications.* Linkoping Studies in Science and Technology. Dissertations—No. 579. Department of Electrical Engineering Linkoping University, SE-581-83 Linkoping, Sweden.

Bergman, N. (2001) Poster Cramer-Rao bounds for sequential estimation. In Doucet, A., Nando de Freitas, and Gordon, N (ed.), *Sequential Monte Carlo methods in practice*. Pp. 321–338. New York: Springer-Verlag.

Dmitriev, S.P., Stepanov O.A. (1998). Nonlinear filtering and navigation. *Proceedings of 5th Saint Petersburg International Conference on Integrated Navigation Systems. Saint Petersburg.* p.138–149.

Doucet, A., Nando de Freitas, and Gordon, N (2001). *Sequential Monte Carlo methods in practice*. New York: Springer-Verlag, 581 p.

Galdos, J.I. (1980). A Cramer-Rao bound for multidimensional discrete-time dynamical systems. *IEEE Trans. Automat. Contr.*, vol. AC-25, N 1, pp. 117–119.

Gustafsson, F. et al. (2002) Particle Filters for Positioning, Navigation and Tracking. *IEEE Transactions on Signal Processing,* vol. 50, N2., pp. 425–437.

Ivanov, V.M., Stepanov, O.A., Korenevski, M.L. (2000). Monte Carlo Methods for a Special Nonlinear Filtering Problem. *Proceedings of 11th IFAC International Workshop Control Applications of Optimization.* vol. 1. – P. 347-353

Koshaev, D.A. (1998). Comparison of lower bounds of accuracy in problems of nonlinear estimation. *Journal of Computer and Systems Sciences International*, vol. 37, N2, pp. 225–229.

Koshaev, D.A. and Stepanov, O.A. (1997). Application of the Rao-Cramer Inequality in Problems of Nonlinear Estimation. *Journal of Computer and Systems Sciences International*, vol. 36, N 2, pp. 220–227.

Ristic B., Arulampalam S., and Gordon, N. (2004). *Beyond the Kalman filter: Particle filters for tracking applications.* Artech House Publishers.

Simandl, M., Kralovec, J., and Tichavsky, P. (2001) Filtering, predictive and smoothing Cramer-Rae bounds for discrete-time nonlinear dynamic systems, *Automatica,* vol. 37, pp. 1703–1716.

Snyder, C., Bengtsson T., Bickel P., and Anderson, B. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, vol. 136, N 12, pp. 4629–4640.

Stepanov, O.A., Vasilyev, V.A., and Dolnakova, A.S. (2013). Cramer-Rao lower bound for parameters of random processes in navigation data processing. *Proceedings of 21st Mediterranean Conference on Control & Automation.* Platanias-Chania, Crete, Greece, June 25–28, pp. 1214–1221.

Stepanov, O.A., Berkovskiy, N.A. (2011). Investigating the calculation error of the optimal Bayesian estimate for a nonlinear problem with the use of the Monte-Carlo method. *Proceedings of 18th World Congress,* Milano, August 28, September 2.

Tichavsky, P., Muravchik, C., Nehorai, A. (1998) Posterior Cramer-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46, 1386–1398.

Van Trees, H. L. (1968). *Detection, estimation, and modulation theory. Part I: Detection, estimation, and linear modulation theory (Part 1).* John Wiley& Sons.

Van Trees, H. L., Bell, K. L. (2007). *Bayesian bounds for parameter estimation and nonlinear filtering/tracking.* Wiley-IEEE Press.

Research supported by the Russian Foundation for Basic Research, project no. *14-08-00347-a.*