Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Investigation of Linear Optimal Estimator

O.A. Stepanov, A.B. Toropov

State Research Center of Russia - Central Scientific & Research Institute Elektropribor
30, Malaya Posadskaya Str., St. Petersburg, 197046, Russia
(Tel. (812) 232-82-53, fax (812) 232-33-76, e-mail: ostepanov@eprib.ru)

Abstract: Features of the linear optimal estimator (LOE) that minimizes the root mean-square (RMS) criterion in the class of linear estimates in nonlinear problems are investigated. The equivalent linear models for measurements have been introduced for the LOE and the Cramer-Rao Bound. The features of the LOE in comparison with the optimal Bayesian estimator, which corresponds to the conditional mean, are studied. Some examples are considered to illustrate the results obtained.

1. INTRODUCTION

A number of estimators, close by their nature, have been developed for nonlinear filtering problems recently. Among them are the so-called unscented Kalman filter, sigma-point Kalman filter, linear regression Kalman filter (Juiler and Uhlmann, 1995, 2004; Li and Jilkov, 2004; Daum, 2005; Wan and Van der Merwe 2001, 2004; Lefebvre, 2005; Stepanov and Amosov, 2006, 2007). In these publications the main attention is centered on computational aspects and advantages of the estimators suggested as compared to other suboptimal estimators, to the extended Kalman filter, for example. At the same time the following fact that is of paramount importance, remains in the background. The above-mentioned estimators are only approximate procedures for calculation of linear optimal estimate, minimizing the RMS criterion only for estimates linearly dependent on measurements. That means that the linear optimal estimate, even accurately determined, is, in its turn, only the suboptimal estimate relative to the Bayesian optimal one that minimizes the RMS criterion without a limitation on the class of estimates used. Thus it is profitable to be able to reveal and evaluate a possible loss of the LOE as compared to the Bayesian optimal estimate, which is the aim of the present paper.

2. PROBLEM STATEMENT

Consider the following problem: to estimate an n-dimensional random vector $\mathbf{x} = [x_1 \ \ldots \ x_n]^{\mathrm{T}}$ by m-dimensional measurements $\mathbf{y} = [y_1 \ \ldots \ y_m]^{\mathrm{T}}$

$$\mathbf{y} = \mathbf{s}(\mathbf{x}) + \mathbf{v}, \tag{1}$$

in which $\mathbf{s}(\mathbf{x}) = [s_1(\mathbf{x}) \ \ldots \ s_m(\mathbf{x})]^{\mathrm{T}}$ is a known m-dimensional function, $\mathbf{v} = [v_1 \ \ldots \ v_m]^{\mathrm{T}}$ is a random vector of the measurement errors. The joint probability density function (p.d.f.) $f(\mathbf{x}, \mathbf{v})$ for the vectors $\mathbf{x}$ and $\mathbf{v}$ is assumed to be known. For simplicity, the random vectors $\mathbf{v}$ and $\mathbf{x}$ are assumed to be independent of each other, i.e. $f(\mathbf{x}, \mathbf{v}) = f(\mathbf{x}) f(\mathbf{v})$, and $\mathbf{v}$ is a zero mean vector. It is known that the Bayesian optimal (in the RMS sense) estimate $\hat{\mathbf{x}}(\mathbf{y})$, minimizing the criterion

$$J = E[(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{y}))^T (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{y}))], \tag{2}$$

and the design error covariance matrix, conditional to the measurements $\mathbf{y}$, are defined as (Meditch, 1969)

$$\hat{\mathbf{x}}^{opt}(\mathbf{y}) = \int \mathbf{x} f(\mathbf{x}/\mathbf{y}) d\mathbf{x}, \tag{3}$$

$$\mathbf{P}^{opt}(\mathbf{y}) = \int (\mathbf{x} - \hat{\mathbf{x}}^{opt}(\mathbf{y}))(\mathbf{x} - \hat{\mathbf{x}}^{opt}(\mathbf{y}))^T f(\mathbf{x}/\mathbf{y}) d\mathbf{x}, \tag{4}$$

where E is the mathematical expectation corresponding to the $f(\mathbf{x}, \mathbf{y})$ and $f(\mathbf{x}/\mathbf{y})$ is the a posteriori p.d.f. of the vector $\mathbf{x}$. In these relations and further integrals are considered to be multiple, with infinite limits. The conditional matrix $\mathbf{P}^{opt}(\mathbf{y})$ characterizes the design estimation accuracy of the state vector for a specified set of measurements $\mathbf{y}$. If the function $\mathbf{s}(\mathbf{x})$ is nonlinear, or $\mathbf{x}$ and $\mathbf{v}$ are not Gaussian, there arises the problem of designing suboptimal estimators (algorithms), economical from the computational viewpoint. Such estimators must provide calculation of both the estimate $\hat{\mathbf{x}}^{sub}(\mathbf{y})$ that is not substantially different in accuracy from the accuracy of the optimal estimate, and the adequate design error covariance matrix $\mathbf{P}^{sub}(\mathbf{y})$. One of the possible variants of designing suboptimal algorithms is based on the approach which provides minimization of criterion (2) in the class of unbiased estimates, linearly dependent on measurements. This estimator will be called a linear optimal estimator (LOE) as opposed to the nonlinear optimal estimator (NOE) in the form of the conditional mean (3).

The aim of this investigation is to study the features of the LOE and to analyze the relationship and dissimilarities between the LOE and NOE.

3. FEATURES OF THE LOE

From the estimation theory it is known that if the mathematical expectations $\overline{\mathbf{x}}$, $\overline{\mathbf{y}}$, the covariance matrices $\mathbf{P}_{\mathbf{xx}}$, $\mathbf{P}_{\mathbf{yy}}$ and the cross covariance matrix $\mathbf{P}_{\mathbf{xy}}$ are specified for the composite vector $\mathbf{x},\mathbf{y}$, then the linear unbiased estimate $\hat{\mathbf{x}}^{lin}(\mathbf{y})$ that minimizes criterion (2) in the class of linear estimates and the design covariance matrix are determined as (Meditch, 1969):

$$\hat{\mathbf{x}}^{lin}(\mathbf{y}) = \overline{\mathbf{x}} + \mathbf{P}_{\mathbf{xy}}\mathbf{P}_{\mathbf{yy}}^{-1}(\mathbf{y}-\overline{\mathbf{y}}) = \overline{\mathbf{x}} + \mathbf{K}^{lin}\left(\mathbf{y}-\overline{\mathbf{y}}\right), \quad (5)$$

$$\mathbf{P}^{lin} = E[(\mathbf{x}-\hat{\mathbf{x}}(\mathbf{y}))(\mathbf{x}-\hat{\mathbf{x}}(\mathbf{y}))^{\mathrm{T}}] = \mathbf{P}_{\mathbf{xx}} - \mathbf{P}_{\mathbf{xy}}\mathbf{P}_{\mathbf{yy}}^{-1}\mathbf{P}_{\mathbf{yx}}. \quad (6)$$

If the function $\mathbf{s}(\mathbf{x})$ is linear, i.e. $\mathbf{s}(\mathbf{x}) = \mathbf{H}\mathbf{x}$, then

$$\overline{\mathbf{y}} = \mathbf{H}\overline{\mathbf{x}}, \ \mathbf{P}_{\mathbf{yy}} = \mathbf{H}\mathbf{P}_{\mathbf{xx}}\mathbf{H}^{\mathrm{T}} + \mathbf{P}_{\mathbf{v}}, \quad \mathbf{P}_{\mathbf{xy}} = \mathbf{P}_{\mathbf{xx}}\mathbf{H}^{\mathrm{T}}, \quad (7)$$

where $\mathbf{P}_{\mathbf{v}}$ is the covariance for $\mathbf{v}$. If the $\mathbf{s}(\mathbf{x})$ is nonlinear, then in finding $\hat{\mathbf{x}}^{lin}$ there arises the problem of calculation of $\overline{\mathbf{y}}$, $\mathbf{P}_{\mathbf{yy}}$ and $\mathbf{P}_{\mathbf{xy}}$. Taking into consideration (1) and the fact that $\overline{\mathbf{v}} = 0$, these values can be determined as

$$\overline{\mathbf{y}} = \int \mathbf{s}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \quad (8)$$

$$\mathbf{P}_{\mathbf{yy}} = \int (\mathbf{s}(\mathbf{x})-\overline{\mathbf{y}})(\mathbf{s}(\mathbf{x})-\overline{\mathbf{y}})^{\mathrm{T}} f(\mathbf{x}) d\mathbf{x} + \mathbf{P}_{\mathbf{v}}, \quad (9)$$

$$\mathbf{P}_{\mathbf{xy}} = \int (\mathbf{x}-\overline{\mathbf{x}})(\mathbf{s}(\mathbf{x})-\overline{\mathbf{y}})^{\mathrm{T}} f(\mathbf{x}) d\mathbf{x}. \quad (10)$$

In order to gain an understanding of the features of the LOE and its relationship with the NOE, let us present the initial measurements (1) by using the linear model of the form

$$\mathbf{y}^{lin} = \overline{\mathbf{y}} + \mathbf{H}^{lin}(\mathbf{x}-\overline{\mathbf{x}}) + \widetilde{\mathbf{v}}, \quad (11)$$

in which $\widetilde{\mathbf{v}}$ is a zero-mean vector with the covariance $\mathbf{P}_{\widetilde{\mathbf{v}}}$. It is assumed that this vector does not correlate with $\mathbf{x}$ and $\mathbf{v}$. Let us find such values of the $\overline{\mathbf{y}}$, $\mathbf{H}^{lin}$ and $\mathbf{P}_{\widetilde{\mathbf{v}}}$ at which the first two moments for the composite vectors $\mathbf{x},\mathbf{y}^{lin}$ and $\mathbf{x},\mathbf{y}$ will be equal. It is clear that in this case the $\overline{\mathbf{y}}$ must coincide with (8), and from Eq. (7) it follows that

$$\mathbf{H}^{lin} = \mathbf{P}_{\mathbf{yx}}\mathbf{P}_{\mathbf{xx}}^{-1}, \quad (12)$$

$$\mathbf{P}_{\widetilde{\mathbf{v}}} = \mathbf{P}_{\mathbf{yy}} - \mathbf{P}_{\mathbf{yx}}\mathbf{P}_{\mathbf{xx}}^{-1}\mathbf{P}_{\mathbf{xy}}, \quad (13)$$

more fully

$$\mathbf{H}^{lin} = \int (\mathbf{s}(\mathbf{x})-\overline{\mathbf{y}})(\mathbf{x}-\overline{\mathbf{x}})^{\mathrm{T}} f(\mathbf{x}) d\mathbf{x}\mathbf{P}_{\mathbf{xx}}^{-1}, \quad (14)$$

$$\mathbf{P}_{\widetilde{\mathbf{v}}} = \mathbf{P}_{ad} + \mathbf{P}_{\mathbf{v}}, \quad (15)$$

where

$$\mathbf{P}_{ad} = \int (\mathbf{s}(\mathbf{x})-\overline{\mathbf{y}})(\mathbf{s}(\mathbf{x})-\overline{\mathbf{y}})^{\mathrm{T}} f(\mathbf{x}) d\mathbf{x} - \mathbf{P}_{\mathbf{yx}}\mathbf{P}_{\mathbf{xx}}^{-1}\mathbf{P}_{\mathbf{xy}}, \quad (16)$$

where $\mathbf{P}_{\mathbf{xy}}$ is calculated by using (10).

Below the representation of measurements in the form of (11) is called 'an Equivalent Optimal Linear Model' (EOLM) for measurements (1). From (15) it follows that the random vector $\widetilde{\mathbf{v}}$ can be presented as a sum of two vectors $\widetilde{\mathbf{v}} = \mathbf{v}_{ad} + \mathbf{v}$, where $\mathbf{v}_{ad}$ is a zero-mean vector with the covariance matrix (16). The vector $\mathbf{v}_{ad}$ does not correlate with $\mathbf{x}$ and $\mathbf{v}$. From the aforesaid it is obvious that the LOE for nonlinear measurements (1) can be developed as a solution of the linear estimation problem for the EOLM defined by (11), (14)-(16).

Thus, based on the EOLM, it is possible to reveal the following special features of the LOE as compared with the NOE.

First of all, the initial nonlinear function is replaced by the linear one, i.e. $\mathbf{s}(\mathbf{x}) \approx \overline{\mathbf{y}} + \mathbf{H}^{lin}(\mathbf{x}-\overline{\mathbf{x}})$. Secondly, to take into account this replacement, a zero-mean, uncorrelated with $\mathbf{x}$ and $\mathbf{v}$, vector of additional methodical errors $\mathbf{v}_{ad}$ is added in the measurement model. Besides, a priori information about $f(\mathbf{x})$ and $f(\mathbf{v})$ for independent vectors $\mathbf{x}$ and $\mathbf{v}$ is replaced with the data about the mathematical expectations and covariances for uncorrelated vectors $\mathbf{x}$ and $\mathbf{v}$. All these features are possible reasons why the LOE will be losing in accuracy in comparison with the NOE. And at last, another specific feature of the LOE in comparison with the NOE is that the design covariance matrix (6) generated in it does not depend on measurements, which means that the unconditional and conditional error covariance matrices coincide.

Note 1. In should be emphasized that the LOE is derived here based on coincidence of the first two moments for the composite vectors $\mathbf{x},\mathbf{y}^{lin}$ and $\mathbf{x},\mathbf{y}$. In other words, there was no aim to choose parameters of the linear measurements (11) in order to minimize, in a certain sense, the difference between the linear and nonlinear functions. However, it is possible to show that the linear approximation $\mathbf{s}(\mathbf{x}) \approx \mathbf{H}\mathbf{x} + \mathbf{c}$ at $\mathbf{H} = \mathbf{H}^{lin}$, $\mathbf{c} = \overline{\mathbf{y}} - \mathbf{H}\overline{\mathbf{x}}$ for the nonlinear function $\mathbf{s}(\mathbf{x})$ minimizes the criterion $J(\mathbf{c},\mathbf{H}) = E\left\{\left[\mathbf{s}(\mathbf{x}) - \mathbf{H}\mathbf{x} - \mathbf{c}\right]^{T}\left[\mathbf{s}(\mathbf{x}) - \mathbf{H}\mathbf{x} - \mathbf{c}\right]\right\}$. Note that a number of suboptimal linear estimators known from literature are designed in accordance with a scheme in which, first, a linear representation for $\mathbf{s}(\mathbf{x})$ is suggested, and then this representation is used for designing a linear estimator. A possible reason why suboptimal estimators differ from the LOE is multiplicity of procedures for finding parameters of linear approximation. Also of great importance is the fact in which manner the additional error is taken into consideration. Bearing in mind the aforesaid, it is appropriate to note that finding the parameters $\mathbf{H}$ and $\mathbf{c}$, reasoning from minimization of the criterion $J(\mathbf{c},\mathbf{H})$, makes

the content of a well known statistical linearization problem. Linear estimators based on the method of statistical linearization were suggested quite a long time ago (Gelb, 1974). Their distinction from the LOE lies in the fact that the estimator presented there does not take account of an additional error caused by replacement of the nonlinear function with its linear analog. This additional error is taken proper account of in the suboptimal linear estimator suggested in (Lefebvre *et al.*, 2005). In fact, these algorithms are similar to the LOE. Their peculiar feature is that the parameters of linear approximation and the covariance matrices of the additional error are calculated with the use of a special computational procedure based on formation of a set of samples $\mathbf{y}^j$ and $\mathbf{x}^j$. Such algorithms are called linear regression Kalman filters.

Note 2. The realization of the LOE reduces to calculation of the first two moments (8)-(10) and subsequent use of Eq. (5), (6). The question of designing economical, from the computational viewpoint, procedures for finding integrals (8)-(10) are not dealt with in this paper. However it should be mentioned that a number of new methods that allow efficient calculation of these moments have been suggested recently. Among them are as follows: the algorithms that use unscented transformation (the so-called unscented and sigma-point Kalman filters) (Juiler *et al.*, 1995, 2004; Van der Merwe and Van., 2004); the algorithms based on the Monte Carlo method and its modifications; the algorithms with the use of neural networks (Stepanov and Amosov, 2006), and the already mentioned linear regression filters (Lefebvre *et al.*, 2005).

4. TECHNIQUE FOR COMPARISON OF LOE AND NOE

The most objective evaluation of the efficiency of the suboptimal estimator in comparison with the NOE is realized by comparing its accuracy with that of the NOE. The comparison should involve the use of the unconditional error covariance matrices defined as

$$\mathbf{G}^{\mu} = \int\int (\mathbf{x} - \hat{\mathbf{x}}^{\mu}(\mathbf{y}))(\mathbf{x} - \hat{\mathbf{x}}^{\mu}(\mathbf{y}))^T f(\mathbf{x},\mathbf{y}) d\mathbf{x} d\mathbf{y}\,, \quad (17)$$

$\hat{\mathbf{x}}^{\mu}(\mathbf{y})$ - estimates obtained by optimal ($\mu = opt$) and suboptimal ($\mu = sub$) estimators. These matrices can be calculated by the method of statistical testing. In particular, for the diagonal elements

$$G^{\mu}_{kk} \approx \frac{1}{L}\sum_{j=1}^{L} ((x_k^j - \hat{x}_k^{\mu}(\mathbf{y}^j))^2,\ k = \overline{1.n}\,, \quad (18)$$

where $\mathbf{x}^j = [x_1^j \ \ldots \ x_n^j]^{\mathrm{T}}$, $\mathbf{y}^j$ are the random vectors simulated according to $f(\mathbf{x})$ and $f(\mathbf{v})$, $j = \overline{1.L}$. The matrix $\mathbf{G}^{opt}$ characterizes the potential accuracy of estimation.

From (18) it follows that to be able to calculate the potential accuracy, one should know the optimal estimate. This estimate can be found, for instance, by the Monte-Carlo method or its modification, aimed at reducing the amount of computations (Stepanov, 1998; Bergman, 1999). Therefore the analysis of efficiency of suboptimal estimators in accordance with the procedure described above involves a considerable amount of simulation. As regards the accuracy analysis, this procedure can be made much simpler by using the Cramer-Rao inequality which allows finding the low bound for the matrix $\mathbf{G}^{opt}$. For the problem under consideration the Cramer-Rao inequality can be written as (Stepanov, 1998; Bergman, 1999)

$$\mathbf{G}^{opt} \geq \mathbf{J}^{-1}\,, \quad (19)$$

where

$$\mathbf{J} = E\left\{\frac{d\ln f(\mathbf{x},\mathbf{y})}{d\mathbf{x}}\left(\frac{d\ln f(\mathbf{x},\mathbf{y})}{d\mathbf{x}}\right)^{\mathrm{T}}\right\}.$$

Let us consider now a particular case of the problem under consideration, introducing the additional assumption about the Gaussian character of the vectors $\mathbf{x}$ and $\mathbf{v}$, then

$$\mathbf{J}^{-1} = \left(\mathbf{P}_{\mathbf{xx}}^{-1} + \int\left[\frac{d\mathbf{s}^{\mathrm{T}}(\mathbf{x})}{d\mathbf{x}}\mathbf{P}_{\mathbf{v}}^{-1}\frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}^{\mathrm{T}}}\right]f(\mathbf{x})d\mathbf{x}\right)^{-1}. \quad (20)$$

The Cramer-Rao Bound (CRB) $\mathbf{J}^{-1}$ characterizes the potentially achievable accuracy. In order to compare the LOE and NOE similarly to the way it was done for the LOE, it is convenient to introduce a 'Linear Equivalent Model' (LEM) for the CRB

$$\mathbf{y}^{lb} = \mathbf{H}^{lb}\mathbf{x}^* + \mathbf{v}^{lb}\,, \quad (21)$$

for which the error covariance matrix for the vector $\mathbf{x}^*$ coincides with the CRB (20). It is not difficult to see that if the vectors $\mathbf{x}^*$ and $\mathbf{v}^{lb}$ are Gaussian with the covariances $\mathbf{P}_{\mathbf{xx}}$, $\mathbf{P}_{\mathbf{v}}^{lb}$, the matrix $\mathbf{J}^{-1}$ will coincide with the error covariance matrix for estimation of the vector $\mathbf{x}^*$ by the measurements (21) in the case that the following equation holds true:

$$\left(\mathbf{H}^{lb}\right)^{\mathrm{T}}\left(\mathbf{P}_{\mathbf{v}}^{lb}\right)^{-1}\mathbf{H}^{lb} = \int\left[\frac{d\mathbf{s}^{\mathrm{T}}(\mathbf{x})}{d\mathbf{x}}\mathbf{P}_{\mathbf{v}}^{-1}\frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}^{\mathrm{T}}}\right]f(\mathbf{x})d\mathbf{x}\,. \quad (22)$$

The choice of $\mathbf{P}_{\mathbf{v}}^{lb}$ and $\mathbf{v}^{lb}$ that provide Equation (22) is not unambiguous, besides calculation of the right part of Eq. (22) can also involve difficulties. To overcome them and to approximately evaluate $\mathbf{J}^{-1}$, it is useful to introduce the following approximate LEM for the CRB

$$\tilde{\mathbf{y}}^{lb} = \tilde{\mathbf{H}}^{lb}\mathbf{x}^* + \mathbf{v}^{lb}\,, \quad (23)$$

where $\mathbf{P}_{\mathbf{v}}^{lb} = \mathbf{P}_{\mathbf{v}}$, $\widetilde{\mathbf{H}}^{lb} = \int \frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}^{\mathrm{T}}} f(\mathbf{x}) d\mathbf{x}$. It is clear that the covariance matrix corresponding to the linear estimation problem of the vector $\mathbf{x}^*$ by the measurement of (23) is determined as

$$\widetilde{\mathbf{P}}^{lb} = \left(\mathbf{P}_{\mathbf{xx}}^{-1} + \int \frac{d\mathbf{s}^{\mathrm{T}}(\mathbf{x})}{d\mathbf{x}} f(\mathbf{x}) d\mathbf{x} \mathbf{P}_{\mathbf{v}}^{-1} \int \frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}^{\mathrm{T}}} f(\mathbf{x}) d\mathbf{x} \right)^{-1}.$$

Using the Cauchy-Bunyakowsky inequality, it is possible to show that $\widetilde{\mathbf{P}}^{lb} \geq \mathbf{J}^{-1}$ (Stepanov, 1998). Thus, $\widetilde{\mathbf{P}}^{lb}$ is the upper estimate for $\mathbf{J}^{-1}$. It is essential that in a number of cases the matrix can be determined in the closed form. Assuming that the potential accuracy determined by $\mathbf{G}^{opt}$ is close to the potentially achievable accuracy specified by $\mathbf{J}^{-1}$ or its estimate $\widetilde{\mathbf{P}}^{lb}$, and taking into consideration the introduced linear equivalent models (11) and (21) or (23), it is possible to compare the NOE with the LOE.

Note 3. An important characteristic in evaluating the efficiency of suboptimal algorithms is adequacy of the design covariance matrix $\mathbf{P}^{sub}(\mathbf{y})$ generated in it. To check the adequacy of $\mathbf{P}^{sub}(\mathbf{y})$, let us use the matrix

$$\widetilde{\mathbf{G}}^{\mu} = \int \mathbf{P}^{\mu}(\mathbf{y}) f(\mathbf{y}) d\mathbf{y} \approx \frac{1}{L} \sum_{j=1}^{L} \mathbf{P}^{\mu}(\mathbf{y}^j). \qquad (24)$$

The design characteristic will be considered adequate if $\mathbf{G}^{\mu}$ coincides with $\widetilde{\mathbf{G}}^{\mu}$. It is clear that for the NOE the design characteristic is always adequate to its real value as $\mathbf{G}^{opt} = \int \mathbf{P}^{opt}(\mathbf{y}) f(\mathbf{y}) d\mathbf{y} = \widetilde{\mathbf{G}}^{opt}$. In the LOE the matrix $\mathbf{P}^{lin}$ is used as a design accuracy characteristic. As follows from (6), the matrix $\mathbf{P}^{lin}$ corresponds to the real value of the unconditional covariance matrix and hence, the design accuracy characteristic generated in the LOE is adequate too.

5. EXAMPLES

EXAMPLE 1. Let us illustrate possible losses in accuracy due to replacement of the nonlinear function with the linear one and presence of additional methodical errors $\mathbf{v}_{ad}$. Consider the following example. Assume that it is required to evaluate a scalar parameter by the scalar measurements of the form

$$y_i = ax + bx^3 + v_i,\ i = \overline{1.m}$$

in which a, b are the known values, and x, v_i – the zero-mean Gaussian values, independent of each other, with the variances σ_0^2 and r^2. It can be shown that here $\overline{\mathbf{y}} = \mathbf{0}$, $\mathbf{P}_{\mathbf{yy}} = \sigma_0^2\left(a^2 + 15b^2\sigma_0^4 + 6ab\sigma_0^2\right)\mathbf{I}_{m\times m} + r^2\mathbf{E}$, $\mathbf{I}_{m\times m}$ - square $m\times m$ matrix of unities, $\mathbf{E}$ - $m\times m$ unity matrix; $\mathbf{P}_{\mathbf{xy}} = \sigma_0^2(a + 3b\sigma_0^2)\mathbf{I}_{1\times m}$, $\mathbf{I}_{1\times m}$ - column $1\times m$ of unities. Thus $\mathbf{H}^{lin} = (a + 3b\sigma_0^2)\mathbf{I}_{m\times 1}$, $\mathbf{P}_{ad} = 6b^2\sigma_0^6\mathbf{I}_{m\times m}$. Therefore, the following representation will hold true for the EOLM in this example

$$y_i^{lin} = h^{lin}x + v_i + v_{ad},\ i = \overline{1.m}, \qquad (25)$$

where $h^{lin} = (a + 3b\sigma_0^2)$, $v_{ad} = d$ is the random bias with the variance $\sigma_d^2 = 6b^2\sigma_0^6$. In other words, in this example the additional methodical error is the systematic component of the error $\widetilde{v}_i$. Taking into consideration (25), it is easy to derive the LOE. Of the greatest interest is the variance for the LOE that can be written as

$$P^{lin} = \left(\sigma_0^{-2} + (\mathbf{H}^{lin})^T \mathbf{P}_{\widetilde{\mathbf{v}}}^{-1} \mathbf{H}^{lin}\right)^{-1} = \frac{\sigma_0^2\left(m\sigma_d^2 + r^2\right)}{\sigma_0^2 m(h^{lin})^2 + m\sigma_d^2 + r^2}. \qquad (26)$$

The CRB can be determined with the LEM

$$y_i^{lb} = h^{lb}x^* + v_i,\ i = \overline{1.m}, \qquad (27)$$

where $h^{lb} = \sqrt{(a^2 + 6ab\sigma_0^2 + 27b^2\sigma_0^4)}$, i.e.

$$J^{-1} = \frac{\sigma_0^2 r^2}{\sigma_0^2 (h^{lb})^2 m + r^2}. \qquad (28)$$

Comparing (25), (26) with (27), (28), it is not difficult to point out the reasons why the LOE can differ in accuracy from the accuracy of the NOE.

Firstly, it may be due to the fact that the coefficient h^{lin} that determines the derivative is lower in the LOE than it is at the CRB calculation. In particular, if the measurement errors is considerably higher than the methodical errors, and it may be assumed that $m\sigma_d^2 << r^2$, from (26) it is possible to derive $P^{lin} \approx \frac{\sigma_0^2 r^2}{\sigma_0^2 m(h^{lin})^2 + r^2}$. Therefore $P^{lin} > J^{-1}$, as $h^{lin} < h^{lb} = \sqrt{(h^{lin})^2 + 18b^2\sigma_0^4}$.

Secondly, the difference is conditioned by an additional error that is represented by a random bias. As the random bias is not observable here, it limits the accuracy for the LOE. If the second summand defining the contribution of the measurement error in (26) can be neglected, then $P^{lin} \approx \frac{\sigma_0^2\sigma_d^2}{\sigma_0^2(h^{lin})^2 + \sigma_d^2}$. In particular, at $\sigma_0^2(h^{lin})^2 >> \sigma_d^2$

and $r^2 \to 0$, it is possible to write $P^{lin} \geq P^* = \frac{\sigma_d^2}{(h^{lin})^2} = \frac{6b^2\sigma_0^6}{(a+3b\sigma_0^2)^2}$.

The approximate LEM for the CRB can be written as

$$\tilde{y}_i = \tilde{h}^{lb} x^* + v_i, \quad i = \overline{1.m},$$

where $\tilde{h}^{lb} = \sqrt{(a^2 + 6ab\sigma_0^2 + 9b^2\sigma_0^4)} = h^{lin}$. In this case the CRB can be determined as

$$\tilde{P}^{lb} = \frac{\sigma_0^2 r^2}{\sigma_0^2 m\left(\tilde{h}^{lb}\right)^2 + r^2}. \qquad (29)$$

Comparing (28) and (29), special attention must be given to the fact that using the upper estimate for the CRB makes it possible to reveal the second, and the most significant, cause for possible limitations in accuracy for the LOE as compared with the NOE.

Below are the results of the simulation, showing a complete support to the conclusions made. Fig.1 presents the graphs of functions $s(x)$ and its linear representations used in the LOE for $\sigma_0 = 1$, $a = 2$, $b = 0.5$. Fig. 2-3 present the plots of $\sqrt{J^{-1}(m)}$, $\sqrt{\tilde{P}^{lb}(m)}$, $\sqrt{P^{lin}(m)}$ as a function of the number of measurements, at $\sigma_0 = 1$, $\sigma_0 = 0.5$, $r = 1$.

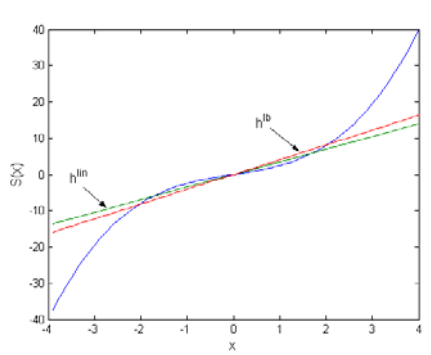

Fig.1 Functions $s(x)$, $h^{lin}x$, $h^{lb}x$.

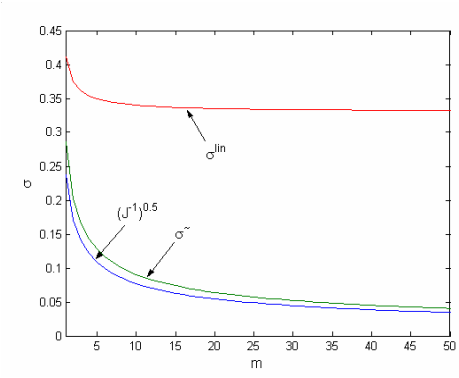

Fig. 2. Plots of $\sqrt{J^{-1}(m)}$, $\sqrt{\tilde{P}^{lb}(m)}$, $\sqrt{P^{lin}(m)}$ at $\sigma_0 = 1$.

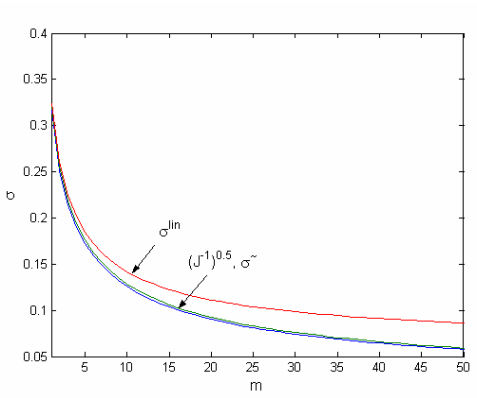

Fig. 3 Plots of $\sqrt{J^{-1}(m)}$, $\sqrt{\tilde{P}^{lb}(m)}$, $\sqrt{P^{lin}(m)}$, $\sigma_0 = 0.5$.

Since, as it was mentioned in the previous section, there is a question about the relationship between the CRB and the real potential accuracy, some more calculations were made in addition to those described above. They show that the potential accuracy in this example, determined by (18) at $\mu = opt$, practically coincides with the CRB. In calculation the optimal estimate was determined with the use of the Monte-Carlo method. Furthermore, some calculations were made that confirm the adequacy of the design accuracy characteristic generated in the LOE. With this end in view, the real error was calculated with the use of (18) at $\mu = lin$, which coincided with the design error, just as it had been expected.

It should be noted that besides the CRB there are some other less conservative bounds (Reece and Nicholson, 2005). The choice of the CRB is conditioned here by a simpler procedure used for its calculation.

From the results given above it follows that the potentialities of the LOE in this example are limited mainly because of the additional error. This drawback of the LOE in the example considered is easy to reveal by comparing the equivalent linear measurements for the LOE and CRB. Using the NOE does not involve any additional error, which is the reason for achieving higher accuracy as compared to the LOE.

EXAMPLE 2. Let us show that the use of a priori information about the mathematical expectations and covariances for uncorrelated vectors $\mathbf{x}$ and $\mathbf{v}$ instead information about $f(\mathbf{x})$ and $f(\mathbf{v})$ can also be a cause of differences in accuracy between the LOE and the NOE. With this aim in view let us assume that for Example 1 $a = 1$, $b = 0$, and x and v_i, $i = \overline{1.m}$ – the uniformly distributed on the interval $[0, 1]$ and $[0, \Delta y]$, correspondingly. In other words, the example of a linear estimation problem is considered, when $\mathbf{x}$ and $\mathbf{v}$ are nonGaussian. It should be noted that the a posteriori p.d.f for such problem will be uniformly distributed on the domain $[c_2, c_1]$. This domain is the crossing of the a priori domain [0,1] and the domain $\Omega \equiv \bigcap_{l=1}^{m} [y_i - \Delta y, y_i] = [y_{\max} - \Delta y, y_{\min}]$,

so $c_1 = \max\{0, y_{\max} - 1\}$, $c_2 = \min\{1, y_{\min}\}$, where $y_{\max}$, $y_{\min}$ - the maximum and minimum values of the measurements. Thus, the NOE and LOE can be determined as

$$\hat{x}^{opt}(y) = \frac{c_2 + c_1}{2}, \qquad P^{opt}(y) = \frac{(c_2 - c_1)^2}{12},$$

$$\hat{x}^{lin}(y) = \frac{1}{2} + \frac{1}{(\Delta y^2 + m)} \sum_{i=1}^{m} \left(y_i - \frac{1 + \Delta y}{2} \right), \; P^{lin} = \frac{\Delta y^2}{12(\Delta y^2 + m)}.$$

The results of $\mathbf{G}^{opt}$ calculation derived using Eq. (18) for the NOE and using P^{lin} for the LOE, when $\Delta y = 0.1$, for various *m* are presented in Table 1.

Table 1. Comparison of $\mathbf{G}^{opt}$ and $\mathbf{G}^{lin} = \mathbf{P}^{lin}$

m	10	20	100
LOE	0.009	0.007	0.003
NOE	0.006	0.003	0.001

From Table 1 it follows that the LOE, even in the linear problem, can substantially lose in accuracy to the NOE.

6. CONCLUSIONS

The equivalent optimal linear models (EOLM) of nonlinear measurements have been introduced for the linear optimal estimator (LOE). Using the EOLM, the features of the LOE and possible reasons why the LOE will be losing in accuracy in comparison with the NOE are investigated.

For the particular case of the problem under consideration when **x** and **v** are Gaussian, the ELM have been introduced for the Cramer-Rao bound. The use of the ELM makes comparison of the LOE and NOE substantially simpler.

Possible losses in accuracy of the LOE in comparison with NOE conditioned by the revealed features are illustrated by examples.

The relation between the LOE, and such estimators as unscented Kalman filter, linear regression Kalman filter and estimators based on the method of statistical linearization have been discussed.

REFERENCES

Bergman, N. (1999). *Recursive Bayesian estimation. Navigation and Tracking Applications*. Linkoping Studies in Science and Technology. Dissertations, 579. Dept. Elect. Eng., Linkoping Univ., SE-581-83 Linkoping, Sweden.

Daum, F. (2005). Nonlinear Filters: Beyond the Kalman Filter. *IEEE Aerospace and Electronic Systems. Tutorials,* **Vol. 20(8)**, pp. 57-71.

Gelb, A. (Editor) (1974). *Applied optimal estimation.* M.I.T. Press, Cambridge, MA.

Juiler, S. J., J. K. Uhlmann and H.F. Durrant-Whyte (1995). A New Approach for Filtering Nonlinear Systems. *In Proc. of American Control Conference*, Seattle, Washington, pp 1628-1632.

Juiler, S. J. and J. K. Uhlmann (2004). Unscented Filtering and Nonlinear Estimation, *Proc. IEEE*, **Vol. 92(3)** , pp 401-422.

Lefebvre, T., H. Bruyninckx and J. De Schutter (2005). *Nonlinear Kalman Filtering for Force-Controlled Robot Tasks*. Springer, Berlin.

Li, X. R. and V.P. Jilkov (2004). A survey of Maneuvering Target Tracking: Approximation Techniques for Nonlinear Filtering. *Proc. 2004 SPIE Conference on Signal and Data Processing of Small Targets*, San Diego, pp 537-535

Meditch, J. S. (1969). *Stochastic optimal linear estimation and control*. Mc. Graw Hill, New York.

Reece, S. and D. Nicholson (2005). Tighter alternatives to the Cramer-Rao lower bound for discrete-time filtering. *Proc. of 7th International Conference on Information Fusion*, **Vol. 1**, pp: 101-106.

Stepanov, O.A. (1998). *Nonlinear filtering and its application in navigation*. State Research Center of Russia - Central Scientific & Research Institute Elektropribor, St. Petersburg, Russia. Pp. 11–343. In Russian.

Stepanov, O.A. and O.S. Amosov (2006). Optimal Estimation Algorithms Based on the Monte Carlo Method and Neural Networks for Nonlinear Navigational Problems. In: *Proc. of the CAC/CACD/ICC,* Munich, Germany.

Stepanov, O. A., O.S. Amosov, A.B. Toropov. (2007). Comparison of Kalman-Type Algorithms in Nonlinear Navigation Problems for Autonomous Vehicles. In: *Proc. of 6-th IFAC Symposium IAV-2007*.

Van der Merwe, R. and E. A. Wan (2001). The Unscented Kalman Filter. In: *Kalman Filtering and Neural Networks* (Haykin S.), pp 221-268, John Wiley & Sons. Inc.

Van der Merwe, R. and E. A. Wan (2004). Sigma-Point Kalman Filters for Integrated Navigation. *ION 60th Annual Meeting U.S. Air Force Institute of Technology*.

This work was supported by the RFBR under Grant 05-08-50189a