An effective recurrence formula for calculating lower Cramer-Rao bounds in case the state-vector is constant

[1]DMITRIY G.ARSENJEV, [2]N.A. BERKOVSKII
[12] St. Petersburg State Polytechnical University,
St. Petersburg, Polytechnicheskaya, 29
RUSSIAN FEDERATION
[1]imop@imop.spbstu.ru, [2]berkovsk@mail.ru, http:// www.spbstu-eng.ru

Abstract: - An effective method for calculating the Bayesian lower unconditional Cramer-Rao bound on condition that the state-vector is constant has been proposed. The recurrence formula for calculating the Fisher information matrix is proved. The method is applicable to arbitrary model noises including non-Gaussian ones. The effectiveness of the approach proposed is shown by applying to the bearing-only tracking problem.

Key-Words Cramer-Rao bound, Bayesian inference, Bearing-only tracking, unconditional covariance matrix, Fisher matrix.

1 Introduction

Bayesian inference is used in signal processing if handling of any uncertainties is required. In Bayesian framework, the potential accuracy of the state-vector estimator is expressed in terms of the unconditional covariance matrix [1]. If the filtering problem solved is nonlinear, that matrix is difficult to evaluate. On the other hand, it is well-known, that the lower bound for the unconditional covariance matrix may be determined using the Cramer-Rao inequality. That bound is called the Cramer-Rao bound (CRB) [2]. The values of CRB depend on only the joint p.d.f. for the measurements and the state-vector, so multiple calculations of the estimator are not necessary. Despite the fact that CRB bound is easier to calculate than the unconditional covariance matrix, one need to evaluate multiple integrals over high-dimensional spaces. At the same time, calculations must be very accurate to guarantee correct computing the Fisher matrix inverse. In essence, the only tool to handle such integrals is the Monte Carlo method, but direct applying is usually ineffective because the high dimensionality entails the large variance of Monte Carlo estimators. Important sampling method [1] decreasing that variance has to be applied to set of integrals, thus one need to seek several important sampling distributions for each element of the Fisher matrix, which results in high computational burden. The remedy to overcome this difficulty is using some recurrence formula, in which expectations would be taken over a low-dimensional space. Such a formula is derived in [2] for the quite general model where the state-vector is varying with time and the measurement and system noises are additive Gaussian. In our paper, we propose a recurrence formula for the special case of the constant state-vector, however, we do not impose any restrictions on the model noises, particularly, the noises are not supposed to be Gaussian. So our formula doesn't follow from one obtained in [2], if the constant state-vector is substituted. It can be shown that our formula and the formula from (2) produce the same results when the noises are additive Gaussian and the state-vector is constant. At the same time, our result doesn't follow from more complicated and general recurrence formula derived in [3]. It should be noted, there is number of the recurrence formulas referred to conditional CRB (see [4, 5]) but in this investigation we are limited by unconditional CRB exclusively. The paper is organized as follows: in section 2 the problem is formulated and denotes are introduced, in the section 3 our recurrence formula is proved, the description of the proposed recurrence algorithm for CRB and the non-recurrence procedure chosen to compare results can be found in section 4. Section 5 is devoted to simulation results, the simplest bearing-only tracking problem is considered, and the conclusion is in the section 6.

2 Problem Formulation

Let the behavior of a discrete-time dynamical system depend on the random vector $\mathbf{v} \in R^n$ with probability density function (p.d.f.) $p(\mathbf{v})$. Suppose the value of $\mathbf{v}$ is unknown, but the noised measurements

$\mathbf{y}_{1:i} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_i]$, $\mathbf{y}_k \in R^m$, $k = \overline{1,i}$ with independent errors are available at each i^{th} moment of time. Let the joint p.d.f. $p(\mathbf{v}, \mathbf{y}_{1:i})$ be known, and our aim is to define at any time t_i $(i = 1,2,\ldots)$ the value of a vector function $\varphi_i(\mathbf{v}): R^n \to R^p$. Let $\hat{\varphi}_i(\mathbf{y}_{1:i})$ be an estimate of $\varphi_i(\mathbf{v})$, then

$$\mathbf{P}_i = \mathsf{E}_{p(\mathbf{v},\mathbf{y}_{1:i})}\left\{(\hat{\varphi}_i(\mathbf{y}_{1:i}) - \varphi_i(\mathbf{v}))(\hat{\varphi}_i(\mathbf{y}_{1:i}) - \varphi_i(\mathbf{v}))^T\right\} \quad (1)$$

is a correlation matrix of estimation errors [5]. Hereafter $\mathsf{E}_{p(.)}$ denotes the expectation w.r.t. p.d.f. $p(.)$. In particular, the diagonal elements of $\mathbf{P}_i$ are mean squared errors of the state-vector $\varphi_i(\mathbf{v})$ estimated components. Then the posterior Cramer Rao lower bound is given by

$$\mathbf{P}_i \geq \mathsf{E}_{p(\mathbf{v})}\left\{\frac{\partial \varphi_i}{\partial \mathbf{v}}\right\} \mathbf{J}_i^{-1} \mathsf{E}_{p(\mathbf{v})}\left\{\frac{\partial \varphi_i}{\partial \mathbf{v}}\right\}^T, \quad (2)$$

where

$$\begin{aligned} \mathbf{J}_i &= \mathsf{E}_{p(\mathbf{v},\mathbf{y}_{1:i})}\{\nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i}))\cdot \\ &\cdot \nabla_{\mathbf{v}}^T \ln(p(\mathbf{v},\mathbf{y}_{1:i}))\} = \\ &= \int \nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i}))\nabla_{\mathbf{v}}{}^T \ln(p(\mathbf{v},\mathbf{y}_{1:i}))\cdot \\ &\cdot p(\mathbf{v},\mathbf{y}_{1:i}) d\mathbf{v} d\mathbf{y}_{1:i} . \end{aligned} \quad (3)$$

The inequality sign in (2) means the difference of left and right matrixes is positive definite. The matrix $\mathbf{J}_i$ is called Fisher information matrix, and $\left\{\frac{\partial \varphi_i}{\partial \mathbf{v}}\right\}$ is the Jacobian matrix of the vector-function φ_i. Product of the matrixes in right side of (2) defines Cramer Rao lower Bound for estimating $\varphi_i(\mathbf{v})$ using $\hat{\varphi}_i(\mathbf{y}_{1:i})$. Inequality (2) holds for the quite wide class of the practically significant estimators $\hat{\varphi}_i(\mathbf{y}_{1:i})$ including biased estimators. The sufficient conditions for holding (2) can be found in [5]. As one can see from (2) and (3), the most difficult task is to evaluate the Fisher matrix, since it is $n+im$-dimensional integral and the calculations may be extremely time-consuming at the large number of measurements i. Next section a recurrence formula for the Fisher matrix is deduced, which allows reducing the problem to the computation of only $n+m$-dimensional integrals at each estimating step.

3 Recurrence formula for the Fisher matrix

Theorem 1. Let the errors of measurements $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_i$ are mutually independent. Suppose the p.d.f. $p(\mathbf{y}_i \mid \mathbf{v})$ is of class C^2 with respect to $\mathbf{v}$. Then the Fisher matrix $\mathbf{J}_i$ in (3) satisfies the recurrence formula

$$\begin{cases} \mathbf{J}_i = \mathbf{J}_{i-1} + \\ + \mathsf{E}_{p(\mathbf{v},\mathbf{y}_i)}\{\nabla_{\mathbf{v}} \ln(p(\mathbf{y}_i \mid \mathbf{v}))\nabla_{\mathbf{v}}^T \ln(p(\mathbf{y}_i \mid \mathbf{v})) - \\ - {}_{\mathbf{v}}^{\mathbf{v}}\Delta p(\mathbf{y}_i \mid \mathbf{v}) / p(\mathbf{y}_i \mid \mathbf{v})\}, \\ \mathbf{J}_1 = \mathsf{E}_{p(\mathbf{v},\mathbf{y}_1)}\left\{\nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_1))\nabla_{\mathbf{v}}^T \ln(p(\mathbf{v},\mathbf{y}_1))\right\} \end{cases} \quad (4)$$

Here, ${}_{\mathbf{v}}^{\mathbf{v}}\Delta p(\mathbf{y}_i \mid \mathbf{v})$ is Hessian of $p(\mathbf{v},\mathbf{y}_1)$ w.r.t. $\mathbf{v}$, that is, ${}_{\mathbf{v}}^{\mathbf{v}}\Delta p(\mathbf{y}_i \mid \mathbf{v}) = \left\{\frac{\partial^2}{\partial v_k \partial v_j} p(\mathbf{y}_i \mid \mathbf{v})\right\}_{k,j=1}^{n}$.

Proof

Since $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_i$ are mutually independent, we have

$$\begin{aligned} p(\mathbf{v},\mathbf{y}_{1:i}) &= p(\mathbf{v})\prod_{k=1}^{i} p(\mathbf{y}_k \mid \mathbf{v}) = \\ &= p(\mathbf{v},\mathbf{y}_{1:i-1})p(\mathbf{y}_i \mid \mathbf{v}), \end{aligned} \quad (5)$$

therefore

$$\begin{aligned} &\ln(p(\mathbf{v},\mathbf{y}_{1:i})) = \ln\left(p(v)\prod_{k=1}^{i} p(\mathbf{y}_i \mid \mathbf{v})\right) = \\ &\ln[p(\mathbf{v},\mathbf{y}_{1:i-1})p(\mathbf{y}_i \mid \mathbf{v})] = \ln(p(\mathbf{v},\mathbf{y}_{1:i-1})) + \\ &+ \ln(p(\mathbf{y}_i \mid \mathbf{v})), \end{aligned}$$

and we get

$$\nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i})) = \nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i-1})) + \nabla_{\mathbf{v}} \ln(p(\mathbf{y}_i \mid \mathbf{v})).$$

Hence,

$$\begin{aligned} &\nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i}))\cdot \nabla_{\mathbf{v}}^T \ln(p(\mathbf{v},\mathbf{y}_{1:i})) = \\ &= [\nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i-1})) + \nabla_{\mathbf{v}} \ln(p(\mathbf{y}_i \mid \mathbf{v}))]\cdot \\ &\cdot [\nabla_{\mathbf{v}}^T \ln(p(\mathbf{v},\mathbf{y}_{1:i-1})) + \nabla_{\mathbf{v}}^T \ln p(\mathrm{p}(\mathbf{y}_i \mid \mathbf{v}))] = \\ &= \nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i-1}))\nabla_{\mathbf{v}}^T \ln(p(\mathbf{v},\mathbf{y}_{1:i-1})) + \\ &+ \nabla_{\mathbf{v}} \ln(p(\mathbf{v},\mathbf{y}_{1:i-1}))\cdot \nabla_{\mathbf{v}}^T \ln(p(\mathbf{y}_i \mid \mathbf{v})) + \\ &+ \nabla_{\mathbf{v}} \ln(p(\mathbf{y}_i \mid \mathbf{v}))\nabla_{\mathbf{v}}^T \ln(p(\mathbf{v},\mathbf{y}_{1:i-1})) + \\ &+ \nabla_{\mathbf{v}} \ln(p(\mathbf{y}_i \mid \mathbf{v}))\nabla_{\mathbf{v}}^T \ln(p(\mathbf{y}_i \mid \mathbf{v})). \end{aligned} \quad (6)$$

Substituting (6) in (3), we obtain another representation of the Fisher matrix:

$$\mathbf{J}_i = \mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\right\}+$$
$$+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}+$$
$$+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\right\}+$$
$$+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\},$$

and using the transposition formula for a matrix product we get equation

$$\begin{aligned}\mathbf{J}_i &= \mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\right\}+\\ &+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}+\\ &+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}^T +\\ &+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}.\end{aligned} \tag{7}$$

Now transform the summands in the right side of (7). Taking into account that

$$\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right) =$$
$$= \left\{\frac{\partial}{\partial v_k}\ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\frac{\partial}{\partial v_j}\ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}_{k,j=1}^{n}, \tag{8}$$

we have the following sequence of equations

$$\mathsf{E}\left\{\frac{\partial}{\partial v_k}\ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\frac{\partial}{\partial v_j}\ln\left(p(\mathbf{v})\right)\right\} =$$
$$= \mathsf{E}\left\{\frac{\frac{\partial}{\partial v_k}p(\mathbf{v},\mathbf{y}_{1:i-1})\frac{\partial}{\partial v_j}p(\mathbf{y}_i \mid \mathbf{v})}{p(v,y_{1:i})}\right\} =$$
$$= \int \frac{\partial}{\partial v_k}p(\mathbf{v},\mathbf{y}_{1:i-1})\frac{\partial}{\partial v_j}p(\mathbf{v})d\mathbf{v}\prod_{l=1}^{i}d\mathbf{y}_l =$$
$$= \int\left[p(\mathbf{v},\mathbf{y}_{1:i-1})\frac{\partial}{\partial v_j}p(\mathbf{y}_i \mid \mathbf{v})\Big|_{v_k=-\infty}^{v_k=\infty} -\right.$$
$$\left.-\int p(\mathbf{v},\mathbf{y}_{1:i-1})\frac{\partial^2}{\partial v_k \partial v_j}p(\mathbf{y}_i \mid \mathbf{v})dv_k\right]d\mathbf{v}_{-k}\prod_{l=1}^{i}d\mathbf{y}_l =$$
$$= -\int p(\mathbf{v},\mathbf{y}_{1:i-1})\frac{\partial^2}{\partial v_k \partial v_j}\mathrm{p}(\mathbf{y}_i \mid \mathbf{v})d\mathbf{v}\prod_{l=1}^{i}d\mathbf{y}_l =$$
$$= -\int\frac{\partial^2}{\partial v_k \partial v_j}p(\mathbf{y}_i \mid \mathbf{v})\left(\int p(\mathbf{v},\mathbf{y}_{1:i-1})\prod_{l=1}^{i-1}d\mathbf{y}_l\right)d\mathbf{v}d\mathbf{y}_i =$$
$$= -\int\frac{\partial^2}{\partial v_k \partial v_j}[p(\mathbf{y}_i \mid \mathbf{v})]p(\mathbf{v})d\mathbf{v}d\mathbf{y}_i =$$
$$= -\mathsf{E}_{p(\mathbf{v},y_i)}\left\{\left[\frac{\partial^2}{\partial v_k \partial v_j}\mathrm{p}(\mathbf{y}_i \mid \mathbf{v})\right]/\mathrm{p}(\mathbf{y}_i \mid \mathbf{v})\right\}. \tag{9}$$

In (9), by $d\boldsymbol{v}_{-k}$ the volume element of the $(n-1)$-dimensional Euclidean space is denoted, that is $d\mathbf{v}_{-k} = \prod_{\substack{s=1\\ s\neq k}}^{n} dv_s$. When integrating by parts in (9), we use the fact that $\lim_{v_k\to\infty} p(\mathbf{v},\mathbf{y}_{1:i-1}) = 0$. Thus, from (8) and (9) we obtain the simplified expression for the sum of the second and third terms in (7):

$$\begin{aligned}&\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}+\\ &+\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\right\}=\\ &=-2\mathsf{E}_{\mathrm{p}(\mathbf{v},\mathbf{y}_i)}\left\{\mathbf{v}\Delta p(\mathbf{y}_i \mid \mathbf{v})/\,p(\mathbf{y}_i \mid \mathbf{v})\right\}.\end{aligned} \tag{10}$$

Then simplify the fourth term in (7):

$$\begin{aligned}&\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}=\\ &=\int\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\cdot\\ &\cdot\left(\int p(\mathbf{v},\mathbf{y}_{1:i-1})\prod_{k=1}^{i-1}dy_k\right)p(\mathbf{y}_i \mid \mathbf{v})dvdy_i =\\ &=\int\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\cdot\\ &\cdot p(\mathbf{v})\,p(\mathbf{y}_i \mid \mathbf{v})dvdy_i =\\ &=\mathsf{E}_{p(\mathbf{v},\mathbf{y}_i)}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{y}_i \mid \mathbf{v})\right)\right\}.\end{aligned} \tag{11}$$

At last, transforming the first term in (7) yields

$$\begin{aligned}&\mathsf{E}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\right\}=\\ &=\int\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\cdot\\ &\cdot p(\mathbf{v},\mathbf{y}_{1:i-1})d\mathbf{v}\prod_{k=1}^{i-1}dy_k\int p(\mathbf{y}_i \mid \mathbf{v})dy_i =\\ &=\mathsf{E}_{p(\mathbf{v},\mathbf{y}_{1:i-1})}\left\{\nabla_{\mathbf{v}} \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\cdot\right.\\ &\left.\cdot\nabla_{\mathbf{v}}^T \ln\left(p(\mathbf{v},\mathbf{y}_{1:i-1})\right)\right\}=\mathbf{J}_{i-1}.\end{aligned} \tag{12}$$

Now the required recurrence relation (4) follows from (10) − (12). Theorem 1 is proved.

The advantage of using (4) in comparison with to direct non-recurrence formula (3) is that the expectation in (3) is taken over the $(im+n)$-dimensional space while in (4) we have to evaluate the i expectations over only $(m+n)$-dimensional space. As numerical experiments have shown, the recurrence formula (4) allows us to speed up the computations significantly. It should be note, the dimension of integrals in the recurrence relations from [2] equals *n*, so if the measurements noise is additive Gaussian, then approach proposed in [2] may be more suitable. However, in case the non-Gaussian noise (for example, a glint noise [6]) the

relationships from [2] can't be applied, while the formula (4) is applicable in that situation.

4 Describing recurrence and non-recurrence algorithms for CRB.

Evaluating CRB, one need to calculate two matrixes $\mathsf{E}_{p(\mathbf{v})}\left\{\frac{\partial\varphi_i}{\partial\mathbf{v}}\right\}$ and $\mathbf{J}_i^{-1}$. The matrix $\mathsf{E}_{p(\mathbf{v})}\left\{\frac{\partial\varphi_i}{\partial\mathbf{v}}\right\}$ can be computed by Monte Carlo method generating elements of Jacobian $\frac{\partial\varphi_i}{\partial\mathbf{v}}(\mathbf{v})$ with p.d.f. $p(\mathbf{v})$. The error of calculating $\mathsf{E}_{p(\mathbf{v})}\left\{\frac{\partial\varphi_i}{\partial\mathbf{v}}\right\}$ may be estimated by the standard deviation value, as it is usually done in Monte Carlo method. As regards the second matrix $\mathbf{J}_i^{-1}$, the situation is more complicated, since the Monte Carlo method has to be applied to $\mathbf{J}_i$ not to $\mathbf{J}_i^{-1}$, and one has to deterrmine the error in calculation of $\mathbf{J}_i^{-1}$ knowing the information about the $\mathbf{J}_i$ calculation error only. For this aim, the Theorem 2 is formulated below without proof.

Theorem 2 [7].

Let $\mathbf{Ax}=\mathbf{f}$ and $(\mathbf{A}+\Delta\mathbf{A})\mathbf{x}=\mathbf{f}+\Delta\mathbf{f}$ be linear systems, $\mathbf{x}\in R^n, \mathbf{f}\in R^n$. Suppose, $\mathbf{A}^{-1}$ exists and $\mu(\mathbf{A})\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}<1$, where $\mu(\mathbf{A})$ is the condition number and $\|\ \|$ stands for compatible vector and matrix norms. Then, if $\mathbf{u}$ is the solution of $\mathbf{Ax}=\mathbf{f}$, and $\mathbf{u}+\Delta\mathbf{u}$ is the solution of $(\mathbf{A}+\Delta\mathbf{A})\mathbf{x}=\mathbf{f}+\Delta\mathbf{f}$ we have

$$\frac{\|\Delta\mathbf{u}\|}{\|\mathbf{u}\|}\le\frac{\mu(\mathbf{A})}{1-\mu(\mathbf{A})\|\Delta\mathbf{A}\|\|\mathbf{A}\|^{-1}}\left(\frac{\|\Delta\mathbf{f}\|}{\|\mathbf{f}\|}+\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}\right) \quad (13)$$

Now let $\Delta\mathbf{J}_i^{-1}$ and $\Delta\mathbf{J}_i$ be the errors of calculating $\mathbf{J}_i^{-1}$ and $\mathbf{J}_i$. It follows from (13) that

$$\left\|\left(\Delta\mathbf{J}_i^{-1}\right)^{\langle k\rangle}\right\| / \left\|\left(\mathbf{J}_i^{-1}\right)^{\langle k\rangle}\right\| \le \frac{\mu(\mathbf{J}_i)\|\Delta\mathbf{J}_i\|}{\|\mathbf{J}_i\|-\mu(\mathbf{J}_i)\|\Delta\mathbf{J}_i\|}; \quad (14)$$

$k=\overline{1,n}$,

where $\left(\Delta\mathbf{J}_i^{-1}\right)^{\langle k\rangle}$ and $\left(\mathbf{J}_i^{-1}\right)^{\langle k\rangle}$ are the k^{th} columns of the matrixes $\Delta\mathbf{J}_i^{-1}$ and $\mathbf{J}_i^{-1}$. Let the inequalities

$$\left\|\left(\Delta\mathbf{J}_i^{-1}\right)^{\langle k\rangle}\right\| / \left\|\left(\mathbf{J}_i^{-1}\right)^{\langle k\rangle}\right\| \le \varepsilon \quad k=\overline{1,n} \quad (15)$$

serve as a criterion for accuracy of approximate computing $\mathbf{J}_i^{-1}$. Then, basing on (13) and (14), the condition for the computational procedure stop may be written as

$$\begin{cases}\mu(\tilde{\mathbf{J}}_i)\|\Delta\tilde{\mathbf{J}}_i\|\left(\|\tilde{\mathbf{J}}_i\|-\mu(\tilde{\mathbf{J}}_i)\|\Delta\tilde{\mathbf{J}}_i\|\right)^{-1}\le\varepsilon\\ \|\tilde{\mathbf{J}}_i\|-\mu(\tilde{\mathbf{J}}_i)\|\Delta\tilde{\mathbf{J}}_i\|>0,\end{cases} \quad (16)$$

where waves are above the approximate values of the corresponding matrixes. Now for the straigtforward non-recurrence method we have relations

$$\begin{cases}\tilde{\mathbf{J}}_i=\sum_{k=1}^{N}\frac{\nabla_{\mathbf{v}}\ln\left(p\left(\mathbf{v}^{(k)},\mathbf{y}^{(k)}{}_{1:i}\right)\right)\cdot\nabla_{\mathbf{v}}^{T}\ln\left(p\left(\mathbf{v}^{(k)},\mathbf{y}^{(k)}{}_{1:i}\right)\right)}{N}\\ \Delta\tilde{\mathbf{J}}_i=3N^{-1/2}\sqrt{\sigma(\tilde{\mathbf{J}}_i)},\end{cases} \quad (17)$$

where $\left(\mathbf{v}^{(k)},\mathbf{y}^{(k)}{}_{1:i}\right)$, $k=\overline{1,N}$ are the random independent vectors with p.d.f. $p\left(\mathbf{v}^{(k)},\mathbf{y}^{(k)}{}_{1:i}\right)$, $\sigma(\tilde{\mathbf{J}}_i)$ is the matrix containing the standard deviations of elements $\tilde{\mathbf{J}}_i$ and N is number of Monte Carlo realizations. Given the accuracy parameter $\varepsilon>0$, the computation stops as soon as the condition (16) is satisfied. When $\mathbf{J}_i$ is calculated using the recurrence formula (4), the Monte Carlo method is applied i times and the algorithm is given by the formulas listed below.

$$\begin{cases}\mathbf{S}_i=\sum_{k=1}^{N}\left(\nabla_{\mathbf{v}}\ln\left(p\left(\mathbf{y}_i{}^{(k)}\mid\mathbf{v}^{(k)}\right)\right)\nabla_{\mathbf{v}}^{T}\ln\left(p\left(\mathbf{y}_i{}^{(k)}\mid\mathbf{v}^{(k)}\right)\right)-\right.\\ \left.-2_{\mathbf{v}}^{\mathbf{v}}\Delta p\left(\mathbf{y}_i{}^{(k)}\mid\mathbf{v}^{(k)}\right)/\,p\left(\mathbf{y}_i{}^{(k)}\mid\mathbf{v}^{(k)}\right)\right)/N, \qquad i>1\\ \tilde{\mathbf{J}}_i=\tilde{\mathbf{J}}_{i-1}+\mathbf{S}_i, \qquad i>1\\ \Delta\mathbf{S}_i=3N^{-1/2}\sqrt{\sigma(\mathbf{S}_i)}, \qquad i>1\\ \tilde{\mathbf{J}}_1=\sum_{k=1}^{N}\frac{\nabla_{\mathbf{v}}\ln\left(p\left(\mathbf{v}^{(k)},\mathbf{y}_1{}^{(k)}\right)\right)\cdot\nabla_{\mathbf{v}}^{T}\ln\left(p\left(\mathbf{v}^{(k)},\mathbf{y}_1{}^{(k)}\right)\right)}{N},\\ \Delta\mathbf{S}_1=3N^{-1/2}\sqrt{\sigma(\tilde{\mathbf{J}}_1)},\end{cases} \quad (18)$$

However the conditions of computing halt in this case are the set of inequalities for each i, namely

$$\begin{cases} \mu(\tilde{\mathbf{J}}_i)\|\Delta\mathbf{S}_i\|\left(\left\|\tilde{\mathbf{J}}_i\right\| - \mu(\tilde{\mathbf{J}}_i)\|\Delta\mathbf{S}_i\|\right)^{-1} \le \varepsilon \\ \left\|\tilde{\mathbf{J}}_i\right\| - \mu(\tilde{\mathbf{J}}_i)\|\Delta\mathbf{S}_i\| > 0 \\ i = 1,2,\ldots \end{cases} \quad (19)$$

In comparison with the non-recurrence procedure (17), some additional error may be accumulated at each step of the recursion (4) when relations (18)-(19) are used but as our numerical experiments have shown, that error is very small and may be ignored.

5 Numerical experiments

To examine the effectiveness of the formula given by (4) we have considered the simplest discrete bearing-only tracking problem [8]. Let the object observed (for example, a ship or submarine) with coordinates $\mathbf{x}(t) = (x^1(t), x^2(t))$ moves according to linear model

$$\mathbf{x}(t_i) = \mathbf{x}_0 + \mathbf{v}t_i$$

where $\mathbf{v} = (v^1, v^2)$ is some constant velocity and $\mathbf{x}_0 = (x_0^1, x_0^2)$ is the initial position of the object.

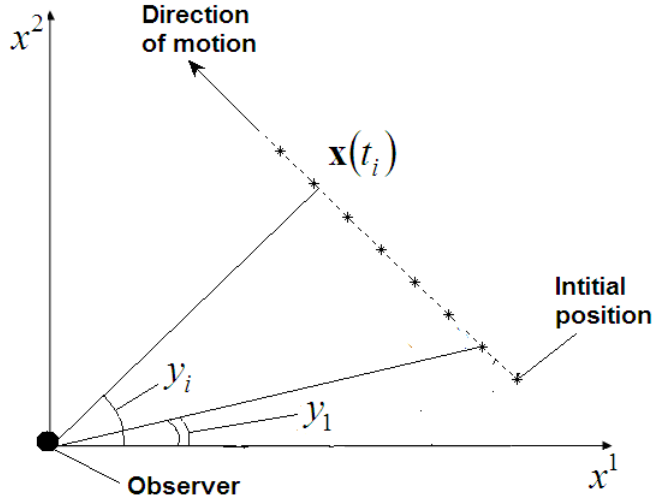

Fig.1 The object's motion and the angles measured.

The initial position $\mathbf{x}_0$ assumed to be known, but the velocity $\mathbf{v}$ is unknown, however, the prior p.d.f. $p(\mathbf{v})$ is available. The observer is located at the origin, and measures the object's bearings y_i at the points of time $t_i = ih$, where h is the time step. These measurements are affected by the random noises, so the model of measurements is defined as follows:

$$y_i = \text{angle}(x_0^2 + v^2 ih, x_0^1 + v^1 ih) + w_i, \quad (20)$$

where

$$\text{angle}(x^1, x^2) = \begin{cases} \tan^{-1}(x^2/x^1), & x^1 \ge 0; \\ \tan^{-1}(x^2/x^1) + \pi, & x^1 < 0,\ x^2 \ge 0; \\ \tan^{-1}(x^2/x^1) - \pi, & x^1 < 0,\ x^2 < 0. \end{cases}$$

The noises w_i, are centered independent identically distributed with the normal p.d.f. $\mathsf{N}(w_i, 0, r) = (2\pi r^2)^{-0,5} \exp(-w_i^2/2r^2)$, $r > 0$. The aim is to determine the object's position $\boldsymbol{x}(t_i)$ at each moment of time t_i on condition that the vector of the measurements $\mathbf{y}_{1:i} = [y_1, y_2, \ldots y_i]$ is available. Let us show that the bearing-only tracking problem is the special case of the more general task described in section 2. Indeed, the motion of the object is fully determined by the unknown velocity vector $\mathbf{v} = (v^1, v^2)$ with the prior p.d.f. $p(\mathbf{v})$, and one is interested in estimating the value of the vector-function $\varphi_i(\mathbf{v}) = \mathbf{x}_0 + \mathbf{v}ih$ given the vector of measurements $\mathbf{y}_{1:i} = [y_1, y_2, \ldots y_i]$ at $t = t_i$. In our simulation, $\mathbf{x}_0 = (955; 100)$, $r = 1^\circ$, $h = 5\text{s.}$, $p(\mathbf{v}) = \mathsf{N}(\mathbf{v}, (-6; 8), 0{,}25\mathbf{E})$, $\varepsilon = 0.1$, where $\mathsf{N}(\mathbf{v}, \mathbf{m}, \Sigma)$ is p.d.f. of the Normal distribution with the expectation $\mathbf{m}$ and the covariance matrix Σ, $\mathbf{E}$ is unitary matrix. Then the joint p.d.f. $p(\mathbf{v}, \mathbf{y}_{1:i})$ in (3) is

$$p(\mathbf{v}, \mathbf{y}_{1:i}) = \mathsf{N}(\mathbf{v}, (-6; 8), 0{,}25\mathbf{E}) \cdot \cdot \prod_{k=1}^{i} \mathsf{N}(y_k, \text{angle}(x_0^2 + v^2 kh, x_0^1 + v^1 kh), r)$$

For the p.d.f. $p(\mathbf{y}_i \mid \mathbf{v})$ and $p(\mathbf{v}, \mathbf{y}_i)$ in the recurrence formula (4) we have

$$p(\mathbf{y}_i \mid \mathbf{v}) = \mathsf{N}(y_i, \text{angle}(x_0^2 + v^2 kh, x_0^1 + v^1 kh), r),$$

and

$$p(\mathbf{v}, \mathbf{y}_i) = \mathsf{N}(\mathbf{v}, (-6; 8), 0{,}25\mathbf{E}) \cdot \cdot \mathsf{N}(y_i, \text{angle}(x_0^2 + v^2 kh, x_0^1 + v^1 kh), r)$$

Since $\varphi_i(\mathbf{v}) = \mathbf{x}_0 + \mathbf{v}ih$ is linear function, then in (2) the matrix $\mathsf{E}_{p(\mathbf{v})}\left\{\frac{\partial \varphi_i}{\partial \mathbf{v}}\right\} = ih\mathbf{E}$. The lover Cramer-Rao bound has been computed both the non-recurrence and recurrence methods. The non-recurrence method uses (3), (16), (17) while the recurrence method employs the recurrence relations

given by (4), (18), (19). The results of modeling are shown in Fig.2. Along with the CRB bounds the conditional potential accuracy lines for the optimal Bayesian estimator [2] are shown. Their values were obtained at the particular value of the measurements' vector $\mathbf{y}_{1:i}$ which was modelled. The Bayesian optimal estimator is defined as

$$\widehat{\varphi}_i^{opt}(\mathbf{y}_{1:i}) = \mathsf{E}_{p(\mathbf{v}|\mathbf{y}_{1:i})}\{\varphi_i(\mathbf{v})\} = \mathsf{E}_{p(\mathbf{v}|\mathbf{y}_{1:i})}\{\mathbf{x}_0 + \mathbf{v}ih\} \quad (21)$$

The potential accuracy of the optimal estimator is described by the conditional covariance matrix

$$\mathbf{P}_i^{opt} = \mathsf{E}_{p(\mathbf{v}|\mathbf{y}_{1:i})}\left\{\left(\widehat{\varphi}_i^{opt}(\mathbf{y}_{1:i}) - \varphi_i(\mathbf{v})\right)\left(\widehat{\varphi}_i^{opt}(\mathbf{y}_{1:i}) - \varphi_i(\mathbf{v})\right)^T\right\} \quad (22)$$

In (21) and (22) by $p(\mathbf{v} \mid \mathbf{y}_{1:i})$ the posterior p.d.f. for the velocity $\mathbf{v}$ is denoted. To calculate $\mathbf{P}_i^{opt}$, we used the non-recurrence important sampling and the Delta-method [9] for estimating the calculation error. The estimate in (21) is optimal in the sense that its mean-square error with respect to the $p(\mathbf{v} \mid \mathbf{y}_{1:i})$ is minimal [10].

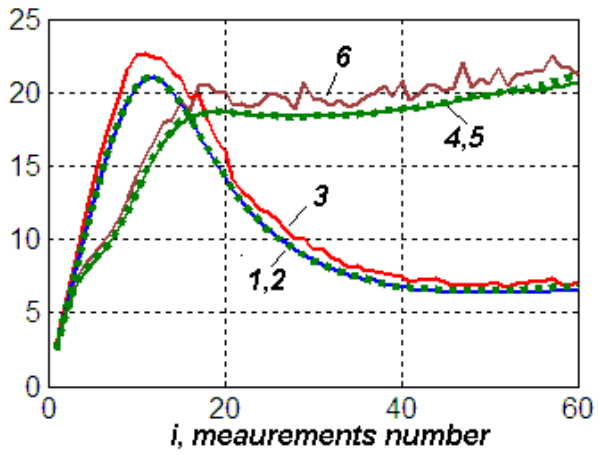

Fig.2 *1,4*– CRBs for x^1 and x^2, recurrence method, *2,5*–CRBs for x^1 and x^2, non-recurrence method, *3,6*– conditional potential accuracy for x^1 and x^2.

Table 1. Run time for both methods.

Number of measure ments, *i*	Run-time, sec.	
	Recurrence method	Non-reccurence method
60	26	1810
30	12	490

In Fig. 2, one can see that CRBs calculated both the recurrence and non-recurrence methods practically coincide, and the potential accuracy lines are above the CBR. Note the distance between CRB and the potential accuracy lines is rather small. From Table 1, it is can be seen that the recurrence method works much faster then the non-recurrence method.

6 Conclusion

A new recurrent formula for unconditional Cramer-Rao bounds has been derived. It is applicable for the constant state-vector without any restrictions imposed on the model noises. Particularly, the non-Gaussian measurements noises are allowed. This formula significantly accelerates the computation of the CRB.

References:

[1] Doucet A. Sequentiual Monte-Carlo methods in practice. /Doucet.A., Freitas N., Gordon N.– N.Y., Shpringer-Verlag, 2001.

[2] D.A. Koshaev, O.A. Stepanov. Application of the Rao-Cramer inequality in problems of nonlinear estimation. Journal of Computer and Systems sciences International 36(2), 1997, 1997, p.220.

[3] P. Tichavsky, C. Muravchik, A Nehorai. Posterior Cramer-Rao bounds for discrete-time nonlinear dynamic system. IEEE Trans. Signal Processing, vol.46, 1998, p.1386.

[4] Z. Long, N. Ruixin, P.K. Varsney, Conditional Posterior Cramer-Rao Lower Bounds for nonlinear recursive Filtering, 12th International Conference on Information fusion, Seattle, WA, USA, Julu 6-9, 2009, p.1528.

[5] N. Bergman. Posterior Cramer-Rao Bounds for Sequential Estimation, in: A. Doucet, N.Freitas, N.Gordon. Sequentiual Monte-Carlo methods in practice. N.Y. Shpringer-Verlag, 2001.

[6] M. A. Masnadi-Shirazi, S. A. Banani. On Maneuvering Target Tracking with Online Observed Colored Glint Noise Parameter //World Academy of Science, Engineering and Technology, 2007, pp. 198 – 201.

[7] Daugavet I. Calculus of approximation theory. Linear equtions. // St. Petersburg, BHV-Petesburg , 2006 [in Russia].

[8] B Ristic, S. Arulampalam, and N. Gordon. Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House, 2004.

[9] N.A. Berkovskii, O.A. Stepanov. Error of calculating the optimal Bayesian estimate using the Monte Carlo method in nonlinear problems, Journal of Computer and Systems sciences International 52(3), 2013 p.342.

[10] O.A. Stepanov, A. Toropov. Investigation of linear optimal estimator. IFAC Proceedings Volumes (IFAC-PapersOnline) 17 (1 PART1), 2008, p.2750.