Distributions of Multiestimates for Statistically Uncertain Systems

Boris I. Anan'ev

***Abstract*— Problems of evolution description are considered for probability distributions (both unconditional and conditional) of random information sets. These sets named in the work as multiestimates naturally appear in problems of state and parameter estimation for statistically uncertain systems. The special attention is given to cases when multiestimates depend on a finite number of parameters. With use of theories of guaranteed and statistical estimation for such cases, the estimation from below is received for the conditional probability of inclusion of a multiestimate in the given set or its intersection with this set. Results are illustrated on examples.**

***Index Terms*— Multiestimates, stochastic inclusions, estimation, filtering.**

I. INTRODUCTION

The concept of information set of a controlled system is successfully used for a long time in the theory of guaranteed estimation of determinate systems, [1]. For systems with the mixed uncertainty including both determinate, and random disturbances, it also were offered estimates in the form of sets, [2], [3]. However, these estimates were not reduced to the information sets in the absence of random disturbances. In this connection, in work [4] the concept of random information set is introduced for multistage systems with the mixed uncertainty. The introduced sets are already reduced to earlier known in the absence of the random disturbances, but generally demand the further processing, as they depend on random and not observable parameters. In work [5], a generalization of the random information sets, named for the brevity multiestimates, is offered for multistage stochastic inclusions. Under some natural assumptions on the inclusions, the multiestimate $X(t, y, \omega)$ at the instant t represents a compact set depending in a not anticipatory way from an element of probability space, formed by stochastic elements, and from the obtained measurements y_0^t. In some sense the multiestimate can be considered as a random state of the statistically uncertain system.

Multiestimates submit to certain evolutionary inclusions. As they are not the equations, but inclusions, we cannot define precisely the probability distributions of the specified elements. However it is possible to define the family of their admissible probability distributions being based on P.Huber's technique [6], and to trace its evolution. Having the admissible family of distributions, it is already enough to allocate simply the conditional distributions when the element y_0^t is fixed, so, to receive the guaranteed estimations from below for probabilities of events of a kind $\{X(t, y, \omega) \subset A\}$ or $\{X(t, y, \omega) \cap A \neq \varnothing\}$, where A is a fixed Borel set. The stated way encounters rather difficult consideration of distributions in the metric space of compact sets, [7]. Therefore in the given work, rather simple special cases of inclusions are considered for which it is possible to reduce the situation to evolution of finite-dimensional random vectors forming a Markov sequence.

This work was supported by the RFBR, grants 04-01-00148, 07-01-00341, and the Fund of Scientific Schools, grant N.Sch. 5344.2006.1.

Institute of Mathematics and Mechanics of UB of RAS, Kovalevskaya str. 16, 620219, GSP-384, Yekaterinburg, Russia. Email: abi@imm.uran.ru

A. *Examples of estimation problems*

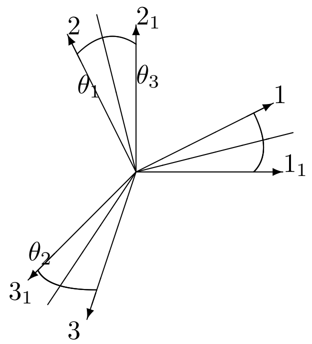
- Let the point of unit mass move on a straight line under the influence of white noise. The movement equations:

$$\ddot{x} = \xi + v, \quad E\xi = 0, \quad \mathrm{cov}(\xi_t, \xi_s) = q\delta(t - s).$$

  The value y is accessible to measurement and submits to the equation

$$\dot{y} = x + \eta + w, \quad E\eta = 0, \quad \mathrm{cov}(\eta_t, \eta_s) = r\delta(t - s).$$

  It is necessary to estimate speed $\dot{x}$ if the initial state x_0 and noises ξ, η are mutually independent. The determinate disturbances v, w have the unknown statistics. Therefore, a speed estimation problem arises with the mixed uncertainty.
- Consider a transport ship-airplane system. Suppose that the base coordinate system (b.c.s.) of the ship is correct. The axis 1 is directed along the parallel to the west. The axis 2 is the local vertical. The axis 3 is directed along the meridian to the north. The position of the airplane (d.c.s.) with respect to b.c.s. is estimated by the Krylov (or Euler) angles.



  The sequence of clockwise rotations: θ_1, θ_3, θ_2. Kinematic equations are of the form

$$\dot{\theta}_1 = \omega_1 - \dot{\theta}_2 \sin\theta_3, \quad \dot{\theta}_2 = (\omega_2 \cos\theta_1$$
$$-\omega_3 \sin\theta_1)/\cos\theta_3, \quad \dot{\theta}_3 = \omega_2 \sin\theta_1 + \omega_3 \cos\theta_1,$$

  where ω_i are projections of relative angular velocity. Under small angles (several degrees) these equations is well linearized:

$$\dot{\theta}_1 = \Omega_1 + \varepsilon_1 - \Omega_{11} + \Omega_{13}\theta_2 - \Omega_2\theta_3,$$
$$\dot{\theta}_2 = \Omega_2 + \varepsilon_2 - \Omega_{12} + \Omega_{11}\theta_3 - \Omega_3\theta_1,$$
$$\dot{\theta}_3 = \Omega_3 + \varepsilon_3 - \Omega_{13} - \Omega_{11}\theta_2 + \Omega_2\theta_1.$$

Here ε_i is the projections of drifts, for which $\dot{\varepsilon}_i = v_i + \xi_i$, $\mid v_i \mid\leq \delta_i$. The values v_i are uncertain, and ξ_i are random. The projections of absolute angular velocity Ω_{1i} are connected with indications a_{1i} of the accelerometers by the relations

$$\dot{\Omega}_{11} = a_{13}/R, \quad \dot{\Omega}_{13} = -a_{11}/R, \quad a_{12} = g - v^2/R,$$
$$v_{11} = -R\Omega_{13}, \quad v_{12} = 0, \quad v_{13} = R\Omega_{11}, \quad \Omega_{12} = 0,$$

where v is the module of velocity of center of mass; R is the radius of the Earth.
The linear approaches of differences of indications of accelerometers in systems d.c.s. and b.c.s. serve as measurements:

$$\dot{y}_1 = a_1 - a_{11} = a_{12}\theta_3 - a_{13}\theta_2 + w_1 + \eta_1,$$
$$\dot{y}_2 = a_2 - a_{12} = -a_{11}\theta_3 + a_{13}\theta_1 + w_2 + \eta_2,$$
$$\dot{y}_3 = a_3 - a_{13} = a_{11}\theta_2 - a_{12}\theta_1 + w_3 + \eta_3,$$
$$\mid w_i \mid\leq \gamma_i,$$

In the simplest case, the movement occurs on equator and $\theta_1 = \theta_2 \equiv 0$. Then only one angle $\theta \equiv \theta_3$ gives a deviation. We have $\dot{\theta} = \Omega - \Omega_1 + \varepsilon$, $\dot{\varepsilon} = v + \xi$; $\Omega \equiv \Omega_3$, $\Omega_1 \equiv \Omega_{13}$. For measurement the difference of outputs of the first accelerometers is used: $\dot{y} = g\theta + w + \eta$.

B. Existing approaches

In 1975 I.Ya. Katz and A.B. Kurzhanski suggest a generalization of Kalman-Busy filter for the problem with mixed uncertainty. Let the inclusions $z(\cdot) = \{v(\cdot), w(\cdot), \hat{x}_0\} \in \mathcal{Z}$ with convex compact constraints $\mathcal{Z}$ be given. It is necessary to find an estimate $\hat{x}^0(t)$ so that

$$\max_{z(\cdot)} E(\|x(t) - \hat{x}^0(t)\|^2 \mid y_0^t)$$
$$= \min_{\hat{x}(t)} \max_{z(\cdot)} E(\|x(t) - \hat{x}(t)\|^2 \mid y_0^t),$$

where $\hat{x}$ is the solution of filter equation with fixed parameters $z(\cdot)$. The estimate $\hat{x}^0(t)$ coincides with Chebyshev center of attainability set $\mathcal{X}(t, y_0^t)$ for filter equation. Later on, I.Ya. Katz have noticed, that if the set $\mathcal{Z}$ is symmetric with respect to $z_0(\cdot)$, then $\dot{\hat{x}}^0 = A\hat{x}^0 + v_0 + PG'R^{-1}(\dot{y} - w_0 - G\hat{x}^0)$, $\hat{x}^0(t_0) = \hat{x}^0_0$. In computing aspect it is equivalent to the Kalman-Busy filter.

Consider a simple example when at once it is more favorable to search for a linear estimate. Given is the scalar equations $\dot{x} = v + \xi$, $\dot{y} = x + w + \eta$, $\int\limits_0^\infty (v^2 + w^2)dt \leq 1$, where $\text{cov}(\xi_t, \xi_s) = q\delta(t - s)$, $\text{cov}(\eta_t, \eta_s) = r\delta(t - s)$; q, $r \geq 0$, $x_0 = 0$. From equation $\dot{p} = q - p^2/r$, $p(0) = 0$, we find $p = \sqrt{qr}\tanh(t\sqrt{q/r})$. The filter:

$$\dot{\hat{x}} = v + p(t)(\dot{y} - w - \hat{x})/r, \; \hat{x}_0 = 0.$$

If $q = 0$, then $\hat{x}^0 \equiv 0$, and maximum of mean-square error equals $\max\limits_v \mid x(t) \mid^2 = t$, i.e. the error grows linearly on time. We will search for an estimate of the form: $f = \int\limits_0^t g(\tau)\dot{y}d\tau + d$. Then it is possible to show, that the optimal estimate satisfies to the equation:

$$\dot{f}^0 = p(t)(\dot{y} - f^0)/(1 + r), \; f^0(0) = 0,$$
$$\dot{p} = 1 + q - p^2/(1 + r), \; p(0) = 0.$$

From this it follows that $p(t) = \tanh(t\sqrt{(1 + q)/(1 + r)}) \times \sqrt{(1 + q)(1 + r)}$ and under $q = 0$ the maximum of mean-square error, equal $p(t)$, is no more than $\sqrt{1 + r}$ for $\forall t$. So, here the linear filter is more preferable.

Let us mention M.L.Lidov and P.E.Eljasberg's early works on linear estimates in problems with the mixed uncertainty. In A.I.Matasov and V.N.Solovyov's works, cases when linear estimates appear optimal among all nonlinear ones are singled out. Different problems with the mixed uncertainty were investigated by the following authors: H.V. Poor, D.P. Loose, J.M. Morris, D.E. Johansen, M. Mintz, I.R. Petersen, A.V. Savkin, G.A.Timofeeva, I.A.Digajlova, A.R. Pankov, K.V.Semenikhin, etc.

II. DEFINITION AND PROPERTIES OF MULTIESTIMATES FOR LINEAR-QUADRATIC SYSTEMS

Consider equations of the form

$$dx = (Ax + Bv)dt + Cd\xi, \quad dy = (Dx + w)dt + d\eta,$$
$$x(0) = x_0, \quad y(0) = 0, \quad x \in R^n, \quad y \in R^m, \quad t \in [0, T]. \tag{II.1}$$

Here $\xi(t)$, $\eta(t)$ are independent Wiener processes with zero means and given covariance matrices for increments, $\text{cov}(d\xi, d\xi) = \text{I}dt$, $\text{cov}(d\eta, d\eta) = \Gamma dt$; $v(t)$, $w(t)$ are determinate disturbances restricted along with initial vector x_0 by the constraints

$$\|x_0\|^2_{P_0} + \int_0^T (\|v(t)\|^2 + \|w(t)\|^2_R) \leq 1, \tag{II.2}$$

where the designation $\|x\|^2_P = x'Px$ is used for symmetric and positively defined matrices P, the symbol $'$ means the transposition. Suppose that $P_0 > 0$, $R > 0$. Then we call an ellipsoid of the form

$$X(t, y, \omega) = \{x : h^2(t) + \|x - x^0(t)\|^2_{P(t)} \leq 1\}, \tag{II.3}$$

where parameters are defined from the equations

$$\dot{P} = -PA - A'P - PBB'P + D'RD, \quad P(0) = P_0,$$
$$dx^0 = Ax^0dt + Cd\xi + P^{-1}D'R(dy - d\eta - Dx^0dt),$$
$$x^0(0) = 0, \quad \dot{h}^2 = \|f_1(t)\|^2_R, \quad h^2(0) = 0, \tag{II.4}$$

a multiestimate $X(t, y, \omega)$. Here the updating function f_1 serves as the unique solution of the equation

$$\int_0^t f_1(\tau)d\tau = y(t) - \eta(t) - \int_0^t Dx^0(\tau)d\tau, \; \forall t \in [0, T]. \tag{II.5}$$

The following statement is true.

Theorem 2.1: For any signal y in system (II.1), (II.2), multiestimate (II.3) is nonempty, i.e. $X(t, y,\ \omega) \neq \emptyset$, and the function f_1 from (II.5) is determined with condition $h^2(T) \leq 1$. Inversely, for any function $f_1 \in L_2^m[0,T]$ with condition $h^2(T) \leq 1$, the signal y of the form

$$\begin{aligned} dy &= d\eta + (Dx^0 + f_1)dt, \quad y(0) = 0, \\ dx^0 &= Ax^0dt + Cd\xi + P^{-1}D'Rf_1dt, \quad x^0(0) = 0, \end{aligned} \tag{II.6}$$

generates nonempty ellipsoid (II.3). The concrete motion $x = x^0 + \delta$, generating signal y, is defined by the equation $\dot{\delta} = (A + BB'P)\delta - P^{-1}D'Rf_1$ with condition $\|\delta(T)\|^2_{P(T)} \leq 1 - h^2(T)$, and by the functions $v(t) \equiv B'P\delta$, $w(t) \equiv f_1 - D\delta$ under condition $x_0 = \delta(0)$.

From theorem 2.1 it follows that only centers x^0 of the ellipsoids are random and, therefore, distributions of multiestimates are defined by the finite-dimensional vectors. From Kalman-Busy theory [11] is known that the conditional distributions of vector $x^0(t)$ will be gaussian, $x^0 \mid y_0^t \sim N(m, \Phi)$, and the parameters $m = \mathrm{E}(x^0 \mid y_0^t)$, $\Phi = \mathrm{cov}(x^0, x^0 \mid y_0^t)$ of the distribution satisfy the system of equations:

$$\begin{aligned} \dot{\Phi} &= A\Phi + \Phi A' + CC' - \Phi D'\Gamma^{-1}D\Phi, \quad \Phi(0) = 0, \\ dm &= (Am + P^{-1}D'Rf_1)dt + \Phi D'\Gamma^{-1}(dy - f_1dt \\ &\quad -Dmdt), \quad m(0) = 0. \end{aligned} \tag{II.7}$$

Let us fix the received results.

Theorem 2.2: For system (II.1), (II.2), the conditional distributions of multiestimates (II.3) are uniquely defined by the family of gaussian distributions for centers x^0, which, in turn, can be found from relations (II.7) under every possible functions f_1 with condition $h^2(T) \leq 1$.

A. Estimation from below for conditional probability of inclusion

Let us write out an estimation from below for conditional probability of inclusion $X(t, y, \omega) \subset K$ for a multiestimate and any convex compact set K. First, the support function of ellipsoid (II.3) looks like $\rho(l \mid X) = l'x^0 + ((1 - h^2)(l'P^{-1}l))^{1/2}$. Second, the event $\{X \subset K\}$ is equivalent to the following $\{\max_{\|l\| \leq 1}\{\rho(l \mid X) - \rho(l \mid K)\} \leq 0\}$, where $\rho(l \mid K)$ is the support function of the compact set K. Therefore, with the account of uncertain parameters, we have the following estimation

$$\min_{f_1(\cdot)} P(\{X(t, y, \omega) \subset K\} \mid y_0^t) = \min_{f_1(\cdot)} \int_{\mathcal{X}_t} f(m, \Phi; x)dx, \tag{II.8}$$

where

$$\begin{aligned} f(m, \Phi; x) &= ((2\pi)^n \det\Phi)^{-1/2} \exp(-\|x - m\|^2_{\Phi^{-1}}/2), \\ \mathcal{X}_t &= \{x \in R^n : \max_{\|l\| \leq 1}\{l'x + ((1 - h^2)(l'P^{-1}l))^{1/2} \\ &\quad -\rho(l \mid K)\} \leq 0\}. \end{aligned} \tag{II.9}$$

We note that the event $\{X \cap K \neq \emptyset\}$ can be expressed via support functions as $\{\min_{x \in K} \sup_{l,q}\{l'x - \rho(l - q \mid X_t) - \rho(q \mid K)\} \leq 0\}$. For this event, also it is possible to write down the estimate of type (II.8).

Suppose that in relations (II.8), (II.9), as a set $K = K_t$, covering the multiestimate at the instant t from above, the ellipsoid

$$\begin{aligned} K_t &= \{x : \|x - k(t)\|^2_{P_t} \leq \varkappa^2\}, \\ dk &= Akdt + \Phi D'\Gamma^{-1}(dy - Dkdt), \ k(0) = 0, \end{aligned} \tag{II.10}$$

is chosen, where $\varkappa > 1$, and centers $k(t)$ are defined by the equation in (II.10) coinciding with the second equation of system (II.7) under $f_1(\cdot) \equiv 0$. The ellipsoid K_t is similar to (II.3). Then estimation (II.8) will not depend on signal realization y_0^t, and we by replacement of variables in integral (II.8) come to the statement.

Theorem 2.3: Let the set $K = K_t$, covering the multiestimate at the instant t from above, has the form (II.10). Then the estimation from below of conditional probability of inclusion $X(t, y, \omega) \subset K_t$ of the form (II.8) is equal

$$\min_{f_1(\cdot)} \int_{\mathcal{Z}_t} f(0, \Phi; x)dx, \tag{II.11}$$

where f is the density of gaussian distribution from (II.9), and $\mathcal{Z}_t$ is the ellipsoid of the form

$$\begin{aligned} \mathcal{Z}_t &= \{x : \|x - z(t)\|^2_{P(t)} \leq (\varkappa - (1 - h^2(t))^{1/2})^2, \\ \dot{z} &= (A - \Phi D'\Gamma^{-1}D)z + (P^{-1}D'R - \Phi D'\Gamma^{-1})f_1, \\ z(0) &= 0. \end{aligned}$$

Owing to theorem 2.3 the estimation from below for probability of inclusion $X(t, y, \omega) \subset K_t$ can be counted up in advance prior to the beginning of estimation process. The nonlinear problem (II.11) of minimization of terminal functional on solutions of the determined system of the equations for $z(t)$ and $h^2(t)$ can be solved by standard methods of optimal control.

B. Example of estimation from below for probability of inclusion

Consider the distribution of multiestimates for two-dimensional system $\dot{x}^1 = x^2$, $\dot{x}^2 = v + \dot{\xi}$ with discrete measurements $y_k = x^1(t_{k-1}) + w_k + \eta_k$, which are spent with constant step $h = T/N$, $t_k = t_{k-1} + h$. Let the uncertain parameters restricted by the constraint

$$p_0((x_0^1)^2 + (x_0^2)^2) + q\int_0^T v^2(t)dt + r\sum_{k=1}^N w_k^2 \leq 1,$$

where constants $p_0, q, r > 0$. The process $\xi(t)$ represents gaussian white noise with covariance of increments $\mathrm{cov}(d\xi, d\xi) = \alpha\, dt$. The values η_k are normally distributed, i.e. $\eta_k \sim N(0, \Gamma)$. We will enter the designations

$$x_k = \begin{bmatrix} x^1(t_k) \\ x^2(t_k) \end{bmatrix}, \quad A = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix},$$

$$v_k = \int_{t_{k-1}}^{t_k} \begin{bmatrix} t_k - \tau \\ 1 \end{bmatrix} v(\tau)d\tau, \quad \xi_k = \int_{t_{k-1}}^{t_k} \begin{bmatrix} t_k - \tau \\ 1 \end{bmatrix} d\xi(\tau).$$

If we are interested in the estimation of states at the moments of measurements' receipt, the initial system is completely equivalent to the discrete one:

$$x_t = Ax_{t-1} + \xi_t + v_t,\ y_t = Dx_{t-1} + w_t + \eta_t,\ t = 1, \ldots, N,$$

where $D = \begin{bmatrix} 1 & 0 \end{bmatrix}$. Here $\xi_t \sim N(0, \Xi)$, where $\Xi = \alpha \begin{bmatrix} h^3/3 & h^2/2 \\ h^2/2 & h \end{bmatrix}$. The restrictions on uncertain parameters are equivalent to the following

$$p_0((x_0^1)^2 + (x_0^2)^2) + \sum_{t=1}^{N} (\|v_t\|_Q^2 + rw_t^2) \leq 1,$$

where $Q = 12q \begin{bmatrix} h^{-3} & -h^{-2}/2 \\ -h^{-2}/2 & h^{-1}/3 \end{bmatrix}$. As a set $K = K_t$ covering the multiestimate at the instant t from above, we choose the ellipse $\{x : \|x - k_t\|_{P_t}^2 \leq \varkappa^2\}$, where $\varkappa > 1$ and the centers k_t are defined from the equation

$$k_t = Ak_{t-1} + A\Phi_{t-1}D'(y_t - Dk_{t-1})/(\Gamma + D\Phi_{t-1}D'),$$

coinciding with the first equation of the system

$$\begin{gathered} m_t = AS_t^{-1}D'R\varphi_t + Am_{t-1} + A\Phi_{t-1}D'(\Gamma \\ +D\Phi_{t-1}D')^{-1}(y_t - \varphi_t - Dm_{t-1}); \\ \Phi_t = A\Phi_{t-1}A' + \Xi \\ -A\Phi_{t-1}D'(\Gamma + D\Phi_{t-1}D')^{-1}D\Phi_{t-1}A', \quad \Phi_0 = 0. \end{gathered}$$

under $\varphi_t \equiv 0$. As well as in a continuous case, the ellipse of the form

$$\{x : \|x - \hat{x}_t\|_{P_t}^2 + \gamma_t \leq 1\}, \tag{II.12}$$

will be a multiestimate here, where parameters satisfy to the equations

$$\begin{gathered} \hat{x}_t = \xi_t + A\hat{x}_{t-1} + AS_t^{-1}D'R\varphi_t, \quad \hat{x}_0 = 0, \\ y_t = \eta_t + D\hat{x}_{t-1} + \varphi_t, \quad \gamma_t = \gamma_{t-1} + \|\varphi_t\|_{T_t}^2, \quad \gamma_0 = 0; \\ P_t^{-1} = AS_t^{-1}A' + BQ^{-1}B', \quad S_t = P_{t-1} + D'RD, \\ P_0 = P_0' > 0, \quad T_t^{-1} = DP_{t-1}^{-1}D' + R^{-1}. \end{gathered} \tag{II.13}$$

In the equations for elements of conditional distribution of a vector $\hat{x}_t$ there are uncertain parameters φ_t which are necessary for considering. The ellipse K_t is similar to (II.12). The signal y_t is modeled according to (II.13) under $\varphi_t = (-1)^t 0.1924$. With a step $h = 0.1$ and parameters $p_0 = q = r = 1/3$ the specified choice φ_t provides the inequality $\gamma_t \leq 1$ for $t = 1, \ldots, N$, where $N = 100$. Let $\alpha = 0.1$, $\Gamma = 0.02$, $\varkappa = 1.3$. We designate the value of of probability's minimum in (II.8) through $Pm(t, y)$. Owing to the set choice K_t, covering the multiestimate, the function $Pm(t, y)$ doesn't depend on a concrete realization of a signal $\{y\}^t$. Its change is shown in Fig. 1. To 30-th step the value of function is stabilized to 0.9111. Covering sets K_t for $t = 5, 10, 15, 20, 25, 30$ under some realization of a signal $\{y\}^t$ are represented in Fig. 2.

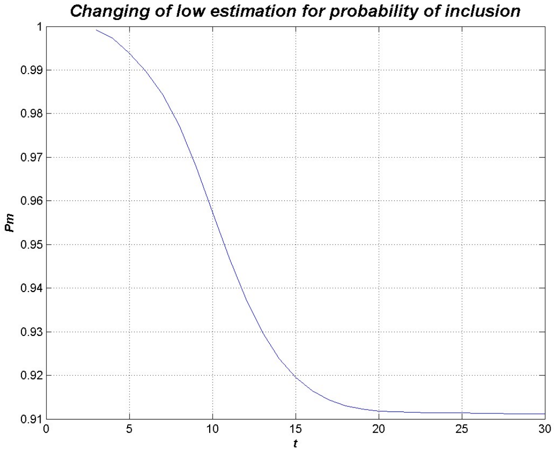


Fig. 1

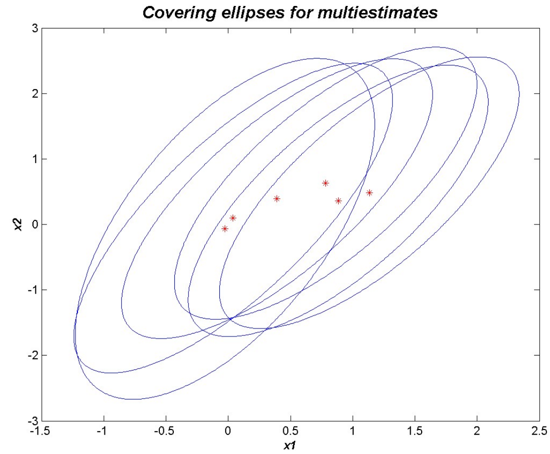


Fig. 2

III. Estimation of interval inclusions

Further on, we write $x \leq y$ if $x^i \leq y^i$, $i = 1, \ldots, n$; $x, y \in R^n$. A set $[x, y] = \{z : x^i \leq z^i \leq y^i, i = 1, \ldots, n\}$ is called an n-dimensional interval (in the sequel, simply interval). Consider interval multistage inclusions of the form

$$x_t \in F(x_{t-1}, \xi_t) \subset R^n, \quad y_t = \{x_{t-1}\}_m + \eta_t \in R^m, \\ t = 1, \ldots, N, \tag{III.1}$$

where $F(x, \xi) = [f^-(x, \xi), f^+(x, \xi)]$ is the interval, y_t is the measured vector, $\{x\}_m$ is the vector composed from the first m coordinates of x, $m \leq n$. The vector functions f^-, f^+ such that $f^-(x, \xi) \leq f^+(x, \xi)$ are supposed to be continuous. The values $\xi_t \in R^k$, $\eta_t \in R^m$ form independent white noise sequences with the known probability distributions. The initial state x_0 is a random vector independent of ξ_t, η_t.

Let $K = [h^-, h^+]$ be an interval in R^n. Then we define the set $\Pi(K, \xi) = [g^-, g^+]$, where $g^{-,i} = \min_{x \in K} f^{-,i}(x, \xi)$, $g^{+,i} = \max_{x \in K} f^{+,i}(x, \xi)$, $i = 1, \ldots, n$. By M^n denote the set of vectors $h = (h^-, h^+) \in R^{2n}$ with property $h^- \leq h^+$. It is clear that the set M^n is homeomorphic to Cartesian product $R^n \times R_+^n$. According to definition of the interval $\Pi(K, \xi)$, there is a mapping taking each vector $h \in M^n$ and each random element ξ to the vector

$$g = G(h, \xi) \in M^n, \qquad \text{(III.2)}$$

where the function $G(h, \xi)$ is continuous owing to the continuity of the functions f^-, f^+ in inclusions (III.1). Note that the obvious inclusion $F(K, \xi) \subset \Pi(K, \xi)$ holds being, as a rule, strong. Here $F(K, \xi)$ is the image of the set K under multivalued mapping F.

Define the following multifunctions

$$\begin{gathered} I(\eta, y) = \{x : \{x\}_m = y - \eta\}, \\ J(K, \xi, \eta, y) = \Pi(K \cap I(\eta, y), \xi), \end{gathered} \qquad \text{(III.3)}$$

and introduce the recurrent relations

$$\begin{gathered} X_t = J(X_{t-1}, \xi_t, \eta_t, y_t), \quad y_t \in \{X_{t-1}\}_m + \eta_t, \\ t = 1, \ldots, N, \end{gathered} \qquad \text{(III.4)}$$

where $X_0 = \{x_0\}$. If we replace the mapping Π by F in relation (III.3), then we get the true multiestimate $X_t^*(y, \xi, \eta)$ as a result of recalculation. Unfortunately, this set cannot be described by finite number of parameters. Introduce its approximation from above by formulas (III.3), (III.4).

Definition 3.1: Any solution of relations (III.4) with some set of elements $\{y\}^t$, $\{\xi\}^t$, $\{\eta\}^t$ is called the approximating multiestimate $X_t(y, \xi, \eta)$ for inclusions (III.1).

The following statement describes some properties of multiestimates.

Lemma 3.1: The solutions of relations (III.4) and corresponding multiestimates have the following properties.

1) Any pair $\{X_t, y_t\}$ serving as a solution of (III.4) consists of the nonempty interval X_t and the vector y_t depending on realizations of random elements $\{\xi\}^t$, $\{\eta\}^t$.
2) The inclusion $X_t^*(y, \xi, \eta) \subset X_t(y, \xi, \eta)$ holds if the sets of elements $\{y\}^t$, $\{\xi\}^t$, $\{\eta\}^t$ are realized in system (III.1).
3) Let the vector $h \in M^n$ specify the set X_{t-1}. Then the multiestimate X_t is given by vector $h_t = G(g_t, \xi_t)$, where $\{g_t^-\}_m = \{g_t^+\}_m = y_t - \eta_t$, and $g_t^{-,i} = h^{-,i}$, $g_t^{+,i} = h^{+,i}$ if $i > m$. Here the function G is defined in (III.2).
4) Due to inclusions in relations (III.4), the random distributions of elements $\{X_t, y_t\}$ remain uncertain.

The basic advantage of the approximating multiestimates consists in the fact that they may be described by finite number of parameters, which can be recalculated due to the simple recurrent formula. More complicated and exact approximations can be obtained, for example, with the help of methods in [8], [9]. However, both the relation of the type (III.2), and further calculation of admissible distributions become also more complicated.

Let us formulate the main problems.

Problem 1: Find recurrent relations for unconditional probability distributions of the sequence $\{X_t, y_t\}$.

Problem 2: Find conditional distributions of the elements X_t and recurrent relations for theirs under the given collection $\{y\}^t$ with the help of the relations of nonlinear filtering.

We can single out separately the parametric case of linear interval inclusions when instead of (III.1) we have the equalities

$$\begin{gathered} x_t = B_t A x_{t-1} + u_t + \xi_t, \quad y_t = \{x_{t-1}\}_m + \eta_t, \\ t = 1, \ldots, N, \end{gathered} \qquad \text{(III.5)}$$

where $B_t = \operatorname{diag}\{v_t^1, \ldots, v_t^n\}$, $v_t \in V$, $u_t \in U$. Here V, U are some intervals. Believing that $F(x, \xi) = \{z : z = BAx + u + \xi, v \in V, u \in U\}$, we can form the sets X_t due to (III.4). However in given case, one can specify the unconditional and conditional distributions for the multiestimates X_t owing to the parametric expression. In particular, one can consider the problem of building of distributions for multiestimates with Gaussian random parameters x_0, ξ_t, η_t.

Note that in the case of diagonal matrices A in equations (III.5), the approximating multiestimates coincide with true ones, i.e. $X_t(y, \xi, \eta) = X_t^*(y, \xi, \eta)$.

Since the second relation in (III.4) is an inclusion, the probability distribution of the pair $\{X_t, y_t\}$ is not exactly defined.

A. Unconditional distributions of multiestimates for interval inclusions

Denote by $\mathcal{P}(M^n)$ the set of all probability measures on Borel σ-algebra $\mathcal{B}(M^n)$ of the space M^n. The space $\mathcal{P}(M^n)$ endowed with the weak topology is a Borel space and, in particular, is metric. Let be given the probability distribution μ of the element ξ and the distribution ν of the element η. For simplicity, we consider that the elements of the sequences ξ_t, η_t have the same distribution. It is convenient to regard the space M^n as a Cartesian product $M^m \times M^{n-m}$, where the space M^m consists of the vectors $(\{g^-\}_m, \{g^+\}_m$ representing the first m coordinates of the components of the vector $g = (g^-, g^+) \in M^n$. Respectively, the space M^{n-m} consists of the last $n-m$ coordinates of the components of the vector g. Then the projection of the vector $g \in M^n$ onto M^m is denoted by $\{g\}_m^M$ and the projection of the same vector onto M^{n-m} is indicated as $[g]_{n-m}^M$.

The evolution of the multiestimates X_t proceeds as follows. Let the vector h_{t-1} correspond the multiestimate at the stage $t-1$ and let $P_{t-1} \in \mathcal{P}(M^n)$ be one of the admissible probability distributions of this vector. At the stage t, the unpredictable choice of the vector

$$q \in [\{h_{t-1}^-\}_m, \{h_{t-1}^+\}_m] \qquad \text{(III.6)}$$

is occurred. Along with the independent realization of the disturbance η_t, this fact gives the realization of observed vector y_t. After that, the vector $g_t = (l, [h_{t-1}]_{n-m}^M) \in M^n$ is allocated, where $l = (q, q) \in M^m$ and $q = y_t - \eta_t$. At last, along with independent realization ξ_t, the nonlinear mapping $h_t = G(g_t, \xi_t)$ is fulfilled. One has to admit the functional dependence $q(h_{t-1})$ for the vector q of the state at the previous stage. Let $q(h_{t-1})$ be a Borel function satisfying inclusions (III.6). It serves as a functional parameter. Let us form the function $\mathcal{G}_q : M^n \times R^k \times R^m \to M^n \times R^m$ by formula

$$\mathcal{G}_q(h,\xi,\eta) = (G(g(h),\xi), q(h)+\eta),$$
$$g(h) = (l(h), [h]_{n-m}^M), \quad l(h) = (q(h), q(h)) \in M^m. \tag{III.7}$$

Doing so the distribution of the pair $\{h_t, y_t\}$ at the stage t can be represented as the Cartesian product of the measures

$$Z_t(B) = (P_{t-1} \times \mu \times \nu)(\mathcal{G}_q^{-1}(B)), \quad B \in \mathcal{B}(M^n \times R^m), \tag{III.8}$$

where P_{t-1} is the admissible distribution of multiestimates at the previous stage, $q(\cdot)$ is the functional parameter.

Generally, formula (III.8) is of interest in connection with the subsequent construction of conditional distributions of the multiestimates. If the distributions of the signals are not necessary, the family $\mathcal{P}_t$ of the unconditional distributions of the multiestimates is constructed similarly to (III.8) by formula

$$\mathcal{P}_t = \bigcup_P \bigcup_{q(\cdot)} (P \times \mu)(G^{-1}), \quad P \in \mathcal{P}_{t-1}, \quad t \geq 2. \tag{III.9}$$

Remark 3.1: In order to simplify the problems, one can suggest the following parametrization: $q(h) = \alpha\{h^+\}_m + (1-\alpha)\{h^-\}_m$, $\alpha \in [0,1]$.

Bearing in mind formulas (III.8), (III.9), we can estimate from below the probability of the events such as $\{X_t \subset K\}$, where $K = [k^-, k^+]$ is some interval. This estimate is of the form $\min_{P\in\mathcal{P}_t} P(\{h_t^- \geq k^-\} \cap \{h_t^+ \leq k^+\})$.

B. Conditional distributions of multiestimates for interval inclusions

Let us use the fact that the values y_t may be observable. Due to [10, th. 7.27] there exists a Borel mapping $r_p(\cdot \mid y) : \mathcal{P}(M^n \times R^m) \times R^m \to \mathcal{P}(M^n)$ such that

$$p(D \times C) = \int_C r_p(D \mid y) p(M^n \times dy) \tag{III.10}$$

for any sets $D \in \mathcal{B}(M^n)$, $C \in \mathcal{B}(R^m)$. The mapping $r_q(\cdot \mid y)$ from (III.10) is fixed. In the same way as in previous section, we construct the single measure $Z_1 \in \mathcal{P}(M^n \times R^m)$ at the stage $t = 1$. Taking into account the measurement y_1, one obtains the conditional distribution $\mathcal{R}_1(y) = \{r_p(\cdot \mid y_1)\}$, $p = Z_1$. Let the family $\mathcal{R}_{t-1}(y) \subset \mathcal{P}(M^n)$, $t \geq 2$, of conditional measures depending of measurements $\{y\}^{t-1}$ be already constructed. Then at the following stage t we have

$$\mathcal{Q}_t(y) = \bigcup_{q(\cdot)} \bigcup \{Z_t : P_{t-1} \in \mathcal{R}_{t-1}(y)\}, \tag{III.11}$$
$$\mathcal{R}_t(y) = \{r_p(\cdot \mid y_t) : p \in \mathcal{Q}_t(y)\}.$$

In formula (III.11), the measures Z_t are implied to be formed due to (III.7), (III.8). The sets $\mathcal{P}_t(y)$ from (III.11) are called the conditional distributions of the multiestimates X_t.

Note that in the special case of single valued right-hand sides in (III.1), i.e, if the equalities hold, the procedure of constructing of the sets (III.11) can be reduced to the well-known procedure of filtering of Markov sequences. In the determinate case, if the sequence $\{y\}^t$ is realized in the system, the conditional distribution of $X_t(y)$ is the uniquely defined information set.

IV. ESTIMATION UNDER COMMUNICATION CONSTRAINTS

Consider only the case described in Section II. Suppose that the estimation is performed via a limited capacity communication channel [12]. In these problems the estimator only observes the transmitted sequence of finite-valued symbols. Therefore, the value $k(t)$ from (II.10) must be coded and later decoded. The channel is able to transmit a code word h at time instants $0, \Delta, 2\Delta, \ldots$ where Δ is a given constant. Introduce the system in R^{2n}:

$$\dot{\tilde{x}}^{(1)} = A\tilde{x}^{(1)}, \quad \dot{\tilde{x}}^{(2)} = (A - KD)\tilde{x}^{(2)} + KD\tilde{x}^{(1)},$$

where $K = \Phi D'\Gamma^{-1}$, on the interval $[i\Delta, (i+1)\Delta)$ with initial conditions $\tilde{x}^{(1)}(i\Delta) = \tilde{x}^{(2)}(i\Delta) = \overline{x}(i\Delta)$, $\overline{x}(0) = 0$.

a) Coder: If $k(i\Delta) - \tilde{x}^{(2)}(i\Delta - 0) \in I_{i_1}^1(a) \times \cdots \times I_{i_n}^n(a) \subset \mathcal{B}_a$, then $h(i) = (i_1, \ldots, i_n)$. Here $I_i^j(a) = \{x^j : x^j \in [-a + 2ai/q, -a + 2a(i+1)/q)\}$, $i = 0, \ldots, q-1$; $\mathcal{B}_a = \{x : \|x\|_\infty \leq a\}$, $\|x\|_\infty = \max_i |x^i|$.

b) Decoder: Let $X(i_1, \ldots, i_n) = (-a + a(2i_1 - 1)/q, \ldots, -a + a(2i_n - 1)/q)$ be the center of the hypercube $I_{i_1}^1(a) \times \cdots \times I_{i_n}^n(a)$. We set $\overline{x}(i\Delta) = X(h(i)) + \tilde{x}^{(2)}(i\Delta - 0)$.

Under some not very hard assumptions one can prove the following: for every $\epsilon > 0$ there exist constants a, Δ and integer q such that $k(i\Delta) - \tilde{x}^{(2)}(i\Delta - 0) \in \mathcal{B}_a$ and $\|k(i\Delta) - \overline{x}(i\Delta)\|_\infty \leq \epsilon$ with probability $\geq 1 - \epsilon$.

REFERENCES

[1] A.B. Kurzhanski, *Control and Estimation under Uncertainty.* Moskow: Nauka, 1977. (in Russian.)
[2] I.Ya. Katz and A.B. Kurzhanski, "Minimax Estimation in Multistage Systems", *Soviet Math. Doklady*, vol. 221, no. 3, pp. 535-538, 1975. (in Russian.)
[3] I.Ya. Katz and A.B. Kurzhanski, "Minimax Multistage Filtering in Statistically Uncertain Situation", *Automatic and Remote Control*, no. 11, pp. 79-87, 1978. (in Russian.)
[4] B.I. Anan'ev, "On Informational Sets for Multistage Statistically Uncertain Systems", *Proceedings of the Steklov Institute of Mathematics, Suppl. 2*, pp. S1-S15, MAIK: 2000.
[5] B.I. Anan'ev, "Distributions of Multi-Estimates for Multistage Stochastic Inclusions", *Journal of Computer and Systems Sciences International*, vol. 46, no. 3, pp. 391-398, 2007. (Transl. from *Izvestiya Akademii Nauk. Teoriya i Sistemy Upravleniya*, no. 3, pp. 59-66, 2007.)
[6] P.J. Huber, *Robust Statistics.* New York: Wiley, 1981.
[7] G. Matheron, *Random Sets and Integral Geometry.* New York: Wiley, 1975.
[8] A.B. Kurzhanski and I. Vályi, *Ellipsoidal Calculus for Estimation and Control.* Boston: Birkhäuser, 1996.
[9] Ye.K. Kostousova, "Inner and Outer Estimation of Attainability Domains for Linear Systems via Parallelotops", *Vychislitelniye Technologii,* vol. 3, no. 2, pp. 11-20, 1998. (in Russian.)
[10] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control.* New York: Academic Press, 1978.
[11] R.Sh. Liptser and A.N. Shiryayev, *Statistics of Random Processes. V.1 General Theory. V.2 Applications.* New York: Springer, 2000.
[12] A.V. Savkin and I.R. Petersen, "Set-Valued State Estimation via a Limited Capacity Communication Channel", *IEEE Trans. on Auto. Control,* vol. 48, no. 4, 2003.