Kybernetika

Ivan Vrana
Quality of optimum statistical estimates in case of ambiguity

QUALITY OF OPTIMUM STATISTICAL ESTIMATES IN CASE OF AMBIGUITY

IVAN VRANA

An ambiguity of measurements is encountered when observations are taken modulo a constant. Paper Vrana [1] has introduced an optimum estimation rule for this case. In the present paper, a behavior of an estimation error is investigated in the case of ambiguous observations. It is shown that the classical approach cannot express the quality of the estimation properly. A three-measure concept is therefore introduced. Equations for these measures and also for bounds are developed.

INTRODUCTION

Paper [1] deals with the problem of the optimum statistical estimation of the quantity X which cannot be observed directly, but whose observation equation has the form

$$F = (X + \Delta) \text{ modulo } \lambda \,, \tag{1}$$

where F is an actual observation, Δ is an observation error and λ is a given constant. Equation (1) is ambiguous because there is generally no one-to-one relation between F and X. Equation (1) is unambiguous only in intervals shorter than λ.

Interval Λ of unambiguous estimation can be considerably extended by performing observations through several different periods λ_i, $i = 1, 2, ..., N$. The quantity Λ is the least common multiple of the used periods λ_i, $\Lambda = n_i \lambda_i$, n_i are mutually non-divisible integers. The global radio navigation system Omega is an example of the technical application of this principle. In the Omega system, the original intervals of unambiguous estimation are given by the wavelengths λ_i of the used signals which are about 30 km. The least common multiple Λ of these wavelengths is about 18 000 km.

The optimum estimation rule of unambiguous estimation was developed in [1] for the case where observation errors are normal random variables. It has been shown there that various criteria of optimality lead to the same form of the optimum estimation rule.

In the present paper we shall investigate the question of how to express the quality

of estimation for the above mentioned case. We shall introduce the concepts of "*trunk*" and "*branches*" of estimation and we shall show that the quality of estimation should be expressed by triple measures: probability of a mistake, the mean value and the variance of a correct estimation.

1. STATISTICAL MEASURES OF QUALITY

In accordance with Vrana [1] we shall use the following notation. Let $\mathbb{R}_N$ be an N-dimensional Euclidean space. The term vector will always mean some element of space $\mathbb{R}_N$. The vectors will be indicated by boldface letters or expressed in a component form, while components are real variables, e.g. $\boldsymbol{D} = (D_1, \ldots, D_N) \in \mathbb{R}_N$.

Analogically as in [1], this notation will also be used

$$\Omega = \underset{i}{\Omega} = \overset{N}{\underset{i=1}{\Omega}}$$

if not indicated otherwise, where the symbol Ω stands for operations $\sum, \prod, \bigcup, \bigcap$.

In [1], we have the following formulation of the estimation problem. We must estimate the real quantity $X \in \langle 0, \varLambda)$ through an observation of the vector $\boldsymbol{F}$ with the components

$$F_i = (X + \varDelta_i) \text{ modulo } \lambda_i \, ; \quad i = 1, \ldots, N \, . \tag{1.1}$$

The value

$$D_i = X + \varDelta_i \tag{1.2}$$

is the ith *true observation*. The true observations are not available to us. Observation errors $\varDelta_i$ are independent zero-mean normal random variables with small variances $\sigma_{\varDelta_i}$. The constants λ_i are mutually commensurable with $\varLambda$ as the least common multiple

$$\lambda_i n_i = \varLambda \, ; \quad i = 1, \ldots, N \, . \tag{1.3}$$

It follows from (1.1) and (1.2) that

$$F_i = D_i \text{ modulo } \lambda_i = d_i(m_i) \text{ modulo } \lambda_i \, ; \quad i = 1, \ldots, N \, . \tag{1.4}$$

A set $\{d_i(m_i)\}$ of so called *fictive observations* d_i is assigned to each D_i

$$d_i(m_i) = D_i + m_i \lambda_i \, ; \quad i = 1, \ldots, N \, , \quad m = \ldots - 1, 0, 1, \ldots \tag{1.5}$$

The fictive observations replace the inaccessible true observations D_i in (1.4). The set of all vectors $\boldsymbol{d} = (d_1, \ldots, d_N)$ will be denoted by $\mathscr{D}$. Every value of the true observation D_i yields a set of fictive observations $\{d_i\}$, according to (1.5). Only one element of this set has the value of the true observation D_i. Analogically, a finite set of vectors of the fictive observations $\{\boldsymbol{d}\}$ corresponds to vector $\boldsymbol{D}$ of the true observations

$$\mathscr{D} := \{\boldsymbol{d}\} \subset \mathbb{R}_N \tag{1.6}$$

and $\boldsymbol{D} \in \mathscr{D}$ also holds.

507

Paper [1] solves the problem of optimum estimation for two typical cases. In the first case, when X is an unknown but nonrandom quantity, the maximum likelihood estimation rule is derived. In the second case, when X is a random quantity with either uniform or normal distribution, the optimum Bayesian estimation rules are developed.

It is shown in [1] that the optimum estimation rule consists of two steps. The first step selects the unique vector $\hat{\boldsymbol{d}}$ from the entire set $\mathscr{D}$ of vectors $\boldsymbol{d}$ and this removes the ambiguity of estimation. The second step then performs the estimation of quantity X using the selected vector $\hat{\boldsymbol{d}}$ of fictive observations. The first step can decide in two ways:

a) $\quad \hat{\boldsymbol{d}} = \boldsymbol{D}$ $\hfill (1.7)$

i.e. the selected vector of fictive observations $\hat{\boldsymbol{d}}$ equals the vector $\boldsymbol{D}$ of the true observations. We shall say that the estimate X belongs to the *trunk* (of estimates) in this case;

b) $\quad \hat{\boldsymbol{d}} \neq \boldsymbol{D}$ $\hfill (1.8)$

which is the opposite case to (1.7). We shall say that the estimate belongs to *branches* (of estimates) here, when the selected vector $\hat{\boldsymbol{d}}$ does not coincide with the true observation $\boldsymbol{D}$.

The estimation from the trunk is the same as the classical optimum estimate (when all true observations are available and, therefore, no ambiguity exists). If $|\Delta_i| < \frac{1}{2}\lambda_i$, then the error of estimation (without a priori information) does not exceed a half of the largest period $\lambda_{i\max}$

$$|\hat{X} - X| < \tfrac{1}{2}\lambda_{i\max} \hfill (1.9)$$

The trunk is the desirable result of the optimum estimation rule. The estimate will be considered as the *correct estimate* if and only if it belongs to the trunk.

On the other hand, if $\hat{X}$ belongs to branches, then the deviation $|\hat{X} - X|$ is generally greater than periods λ_i or observation errors Δ_i. The branches are undesirable results of the optimum estimation rule. If the estimate belongs to the branches, then this event will be considered as a *mistaken estimate*. It is possible to use some suitable method (e.g. coincidence) to distinguish and remove mistakes when repeated estimations are performed.

In our case, the quality of estimation shall be expressed by the following three measures:

1) Probability of the correct estimate P_c;
2) Average value of the correct estimate $E_c = \mathsf{E}\{\hat{X} \mid \hat{\boldsymbol{d}} = \boldsymbol{D}\}$;
3) Variance of the correct estimate $\sigma_c^2 = \sigma^2\{\hat{X} \mid \hat{\boldsymbol{d}} = \boldsymbol{D}\}$.

We have seen in [1] that:

— estimates of a nonrandom quantity X with a uniformly distributed additional observations ϑ, or without any additional observation and also the estimates of a random quantity X with a uniform distribution $p(X)$ have an identical structure

of the optimum estimation rule consisting of the following two steps:

1) Find the vector $\hat{\boldsymbol{d}} \in \mathscr{D}$ for which it holds that

$$\hat{\boldsymbol{d}}: \min_{\boldsymbol{d} \in \mathscr{D}} \sum_{i=1}^{N-1} \sum_{n=i+1}^{N} [(d_i - d_n)/\sigma_{\Delta_i}\sigma_{\Delta_n}]^2 := \sum_{i=1}^{N-1} \sum_{n=i+1}^{N} [(\hat{d}_i - \hat{d}_n)/\sigma_{\Delta_i}\sigma_{\Delta_i}]^2 \quad (1.10)$$

2) Calculate

$$\hat{X} = (\sum_{i=1}^{N} \hat{d}_i/\sigma_{\Delta_i}^2)/a \quad (1.11)$$

where

$$a = \sum 1/\sigma_{\Delta_i}^2 = \text{const} . \quad (1.12)$$

We shall use for all these cases the common name: *estimates without a priori information*;

— estimates of a nonrandom quantity with a Gaussian additional observation ϑ and also the estimates of a Gaussian random quantity X have an identical optimum estimation rule consisting also of two steps:

1) Find the vector $\hat{\boldsymbol{d}} \in \mathscr{D}$

$$\hat{\boldsymbol{d}}: \min_{\boldsymbol{d} \in \mathscr{D}} \left\{ \sum_{i=1}^{N-1} \sum_{n=i+1}^{N} [(d_i - d_n)/\sigma_{\Delta_i}\sigma_{\Delta_n}]^2 + \sum [(d_i - \vartheta)/\sigma_{\Delta_i}\sigma_{\vartheta}]^2 \right\} :=$$

$$:= \sum_{i=1}^{N-1} \sum_{n=i+1}^{N} [(\hat{d}_i - \hat{d}_n)/\sigma_{\Delta_i}\sigma_{\Delta_n}]^2 + \sum [(\hat{d}_i - \vartheta)/\sigma_{\Delta_i}\sigma_{\vartheta}]^2 . \quad (1.13)$$

2) Calculate

$$\hat{X} = (\vartheta/\sigma_{\vartheta}^2 + \sum \hat{d}_i/\sigma_{\Delta_i}^2)/(1/\sigma_{\vartheta}^2 + a) . \quad (1.14)$$

We shall call these cases the *estimate with a priori information*.

2. QUALITY OF ESTIMATE WITHOUT A PRIORI INFORMATION

The optimum estimation rule (1.10) and (1.11) can be expressed also in the form: For each vector $\boldsymbol{d} \in \mathscr{D}$ calculate a pair of quantities

$$\tilde{X} = (\sum d_i/\sigma_{\Delta_i}^2)/a = \tilde{X}(\boldsymbol{d}) \quad (2.1)$$

$$h = \sum (\tilde{X} - d_i)^2/\sigma_{\Delta_i}^2 = h(\boldsymbol{d}) . \quad (2.2)$$

That value $\tilde{X} \in \{\tilde{X}(\boldsymbol{d})\}$ is the optimum estimate $\hat{X}$ to which the smallest value of $h(\hat{\boldsymbol{d}})$ belongs; $\boldsymbol{d} \in \mathscr{D}$.

The fictive observations can be expressed in the form

$$d_i = D_i + m_i\lambda_i ; \quad i = 1, ..., N , \quad m_i \text{ is an integer} . \quad (2.3)$$

Substituting $d_i = D_i$ into (2.1) we obtain a pair of values $\tilde{X} = \tilde{X}_T$ and $h = h_T$ which correspond to the trunk

$$\tilde{X}_T = (\sum D_i/\sigma_{\Delta_i}^2)/a \quad (2.4)$$

$$h_T = \sum (\tilde{X}_T - D_i)^2/\sigma_{\Delta_i}^2 . \quad (2.5)$$

Analogically, substituting (2.3) into (2.1) and (2.2) we get, for each $d \neq D$, the pair $\tilde{X}_B$ and h_B which corresponds to branches.

$$\tilde{X}_B = \tilde{X}_T + \left(\sum m_i \lambda_i / \sigma_{\Delta_i}^2\right)/a \; ; \quad m \neq 0 = \{0, \dots, 0\} \, , \tag{2.6}$$

$$h_B = h_T + Y + Z \tag{2.7}$$

where

$$Y = \sum \alpha_i \delta_i \tag{2.8}$$

$$\alpha_i = 2\left[m_i \lambda_i - \left(\sum_j m_j \lambda_j / \sigma_{\Delta_j}^2\right)/a\right]/\sigma_{\Delta_i} = \text{const} \, . \tag{2.9}$$

$$Z = \sum (\alpha_i/2)^2 = \text{const.} \tag{2.10}$$

$$\delta_i = \Delta_i / \sigma_{\Delta_i} \, . \tag{2.11}$$

The normalized observation errors δ_i, $i = 1, \dots, N$, are independent random variables and have the Gaussian distribution $\mathcal{N}(0, 1)$. The quantities $Z = Z(m)$, $Y = Y(m)$, $\alpha_i = \alpha_i(m)$ are the functions of the vector m.

We get the correct estimate if and only if it holds

$$h_B > h_T \tag{2.12}$$

for all vectors $m \neq 0$. Then it should hold

$$Y(m) > -Z(m) \tag{2.13}$$

according to (2.7).

Let $\mathcal{M}$ be a set of all vectors m for which it holds: $d_i \in \langle D_i - \tfrac{1}{2}\Lambda; D_i + \tfrac{1}{2}\Lambda \rangle$; $d \neq D$. A complementary vector $m^* \in \mathcal{M}$ exists to every $m \in \mathcal{M}$ and it has components

$$m_i^* = -m_i \, . \tag{2.14}$$

Substituting this into (2.8), (2.9) and (2.10) we get

$$\alpha_i(m^*) = -\alpha_i(m) \, , \tag{2.15}$$

$$Y(m^*) = -Y(m) \, , \tag{2.16}$$

$$Z(m^*) = Z(m) \, . \tag{2.17}$$

Vector m^* does not cause, a mistaken estimate if it holds

$$Y(m^*) > -Z(m^*) \, . \tag{2.18}$$

Using (2.16) and (2.17), the last relation (2.18) can be arranged into the form

$$Y(m) < Z(m) \, . \tag{2.19}$$

According to (2.13) and (2.19), the mistake does not happen for the complementary pair of vectors m, m^*, if

$$-Z(m) < Y(m) < Z(m) \, . \tag{2.20}$$

Remark. As we have seen in [1], the process of estimation has a periodic solution with a period Λ. We can, therefore, follow through this process for fictive observations

from any interval of the length Λ, e.g. $d_i \in \langle c; c + \Lambda \rangle$. Choosing this interval symmetrically around D_i, we gain a simple expression of the complementary vector $\boldsymbol{m}^*$ by (2.14). The general position of the interval gives a more complex expression for the complementary vector $\boldsymbol{m}^*$ but finally we obtain the result (2.10) again.

For the given vector of normalized observation errors $\boldsymbol{\delta}$, we obtain the correct estimate if and only if the mistake does not happen for any complementary pair of vectors $\boldsymbol{m}$, $\boldsymbol{m}^* \in \mathcal{M}$, i.e. if it holds

$$Y(\boldsymbol{m}) \in (-Z(\boldsymbol{m}); Z(\boldsymbol{m})) ; \quad \boldsymbol{m} \in \mathcal{M} . \tag{2.21}$$

Let us consider $\boldsymbol{\delta}$ as the vector of coordinates of a point in the N-dimensional space $\mathbb{R}_N$. Then (2.21) defines a volume $\mathscr{V}(\boldsymbol{m})$ between two parallel *hyperplanes* $\mathscr{U}(\boldsymbol{m})$, $\mathscr{U}(\boldsymbol{m}^*)$ for each $\boldsymbol{m} \in \mathcal{M}$. This pair of hyperplanes is defined as follows

$$\mathscr{U}(\boldsymbol{m}), \mathscr{U}(\boldsymbol{m}^*): \sum \delta_i \alpha_i(\boldsymbol{m}) = \pm Z(\boldsymbol{m}) . \tag{2.22}$$

The equation defining $\mathscr{U}(\boldsymbol{m})$ differs from the one for $\mathscr{U}(\boldsymbol{m}^*)$ only by a different sign on the right hand side. The volume $\mathscr{V}(\boldsymbol{m})$ between these hyperplanes will be called the *hyperlayer*. Both hyperplanes, $\mathscr{U}(\boldsymbol{m})$ and $\mathscr{U}(\boldsymbol{m}^*)$, have the same distance r from the origin

$$r = Z/\sqrt{\left(\sum \alpha_i^2\right)} = \tfrac{1}{4} \sqrt{\left(\sum \alpha_i^2\right)} = r(\boldsymbol{m}) . \tag{2.23}$$

If, for instance, $N = 2$, then (2.22) are equations of two parallel straight lines $\delta_1 \alpha_1 + \delta_2 \alpha_2 = \pm Z$ and the hyperlayer $\mathscr{V}(\boldsymbol{m})$ is a strip of the plane between these two parallel straight lines. If $N = 3$ then (2.22) are equations of two parallel planes $\delta_1 \alpha_1 + \delta_2 \alpha_2 + \delta_3 \alpha_3 = \pm Z$ and the hyperlayer $\mathscr{V}(\boldsymbol{m})$ is a part of a 3-D space between these two parallel planes, etc. An example of this situation is shown in Fig. 1 for $N = 2$ and one complementary pair $\boldsymbol{m}$, $\boldsymbol{m}^* \in \mathcal{M}$.
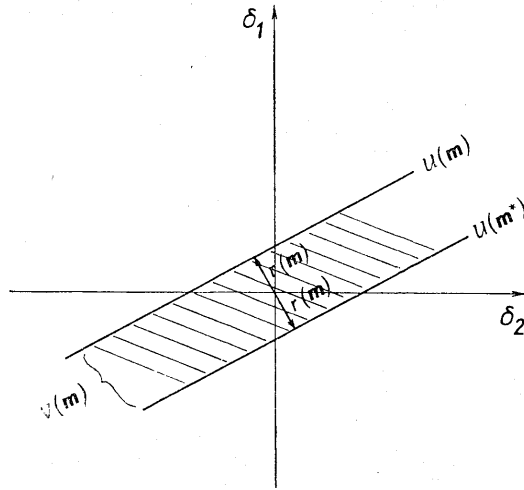

Fig. 1. Hyperplanes, hyperlayer and its width in space $\mathbb{R}_2$.

The equation of the correct estimate (2.21) is equivalent to the following relationship

$$\delta \in \bigcap_{m \in \mathcal{M}} \mathcal{V}(m) := \mathcal{W} .$$ (2.24)

The correct estimate is therefore obtained if and only if the point with the coordinates δ falls into the volume $\mathcal{W}$. The probability of the correct estimate P_c is then

$$P_c = \int \dots \int_{\mathcal{W}} p(\delta) \, \mathrm{d}\delta .$$ (2.25)

In a general case, it might be very difficult to calculate P_c from (2.25). In addition to the difficulties of calculation of the integral itself, difficulties exist already with an analytical expression of the integration domain $\mathcal{W}$. We have, in fact, two possibilities to determine the probability of the correct estimate:

1) To use a suitable Monte-Carlo simulation method. For instance, generate the vectors δ with independent Gaussian components and test (2.21) after each attempt;

2) To find an upper bound P_{cU} and a lower bound P_{cL} for the probability of a correct estimate P_c. We gain the upper bound e.g. by integration in (2.25) over the domain $\mathcal{W}_U$ where $\mathcal{W} \subset \mathcal{W}_U$. Similarly, we get the lower bound by integration over the domain $\mathcal{W}_L \subset \mathcal{W}$. It just remains to make suitable selection for domains $\mathcal{W}_U$, $\mathcal{W}_L$ so that it will be possible to calculate the corresponding integrals.

Use of the simulation method is straightforward. In the following two sections we shall deal with the problems of calculating P_{cU} and P_{cL}.

2.1. Upper Bound

Let us arrange the hyperlayers $\mathcal{V}(m)$ in accordance with an increasing order of their "widths" $2r(m)$ and denote them $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \dots$. The thinnest one is the hyperlayer $\mathcal{V}_1$ which is determined by a pair of hyperplanes $\mathcal{U}_1$ and its width is $2r_1$, etc. Clearly it holds

$$\mathcal{W} \subset \mathcal{V}_1 .$$ (2.26)

Individual pairs of hyperplanes $\mathcal{U}_j$ can have general directions with respect to axes of used coordinate system $\mathbb{C}$. Let us create a new coordinate system $\mathbb{C}'$ by a rotation of $\mathbb{C}$ around the origin so as to obtain the first axis of $\mathbb{C}'$ perpendicular to the hyperplanes $\mathcal{U}_1$. Some functional relation

$$\delta_i = g_i(\delta') ; \quad i = 1, \dots, N$$ (2.27)

holds between the coordinates δ in $\mathbb{C}$ and coordinates δ' in $\mathbb{C}'$. It holds

$$\int \dots \int_{\mathcal{W}} p(\delta) \, \mathrm{d}\delta = \int \dots \int_{\mathcal{W}'} p[g_1(\delta'), \dots, g_N(\delta')] \, |J| \, \mathrm{d}\delta'$$ (2.28)

where J is Jacobian of the transformation (2.27). The transformation $\mathbb{C} \to \mathbb{C}'$ is orthogonal, see Dettman [2, §1.7] and thus $|J| = 1$ and

$$\sum \delta_i^2 = \sum \delta_i'^2 .$$ (2.29)

The joint p.d.f. of the normalized observation error (2.11) has the form

$$p(\boldsymbol{\delta}) = \left(1/\sqrt{(2\pi)}\right)^N \exp\left(-\textstyle\sum \delta_i^2/2\right). \tag{2.30}$$

According to (2.29)

$$p(\boldsymbol{\delta}') = \left(1/\sqrt{(2\pi)}\right)^N \exp\left(-\textstyle\sum \delta_i'^2\right). \tag{2.31}$$

We can see that the coordinates δ_i' are statistically independent Gaussian random quantities $\mathcal{N}(0, 1)$.

In the coordinate system $\mathbb{C}'$, the hyperplanes $\mathcal{U}_1$ are generally given by

$$\mathcal{U}_1\colon \textstyle\sum \delta_i' \alpha_{i,1}' = \pm Z_1'. \tag{2.32}$$

It holds

$$\alpha_{i,1}' = 0 \quad \text{for} \quad i = 2, \ldots, N \tag{2.33}$$

due to perpendicularity of the coordinate axis δ_1' to the hyperplanes $\mathcal{U}_1$. Thus

$$\mathcal{U}_1\colon \delta_1' \alpha_{1,1}' = \pm Z_1'. \tag{2.34}$$

Then, according to (2.23)

$$\mathcal{U}_1\colon \delta_1' = \pm r_1. \tag{2.35}$$

Therefore, the hyperlayer $\mathcal{V}_1$ is in $\mathbb{C}'$ given by

$$\mathcal{V}_1\colon \delta_1' \in (-r_1, r_1). \tag{2.36}$$

Clearly it holds $\mathcal{W} \subset \mathcal{V}_1$ and we can put

$$\mathcal{W}_U = \mathcal{V}_1. \tag{2.37}$$

Now we obtain a very simple equation

$$P_{cU} = \int_{-r_1}^{r_1} p(\delta_1') \, \mathrm{d}\delta_1' = \left(1/\sqrt{(2\pi)}\right) \int_{-r_1}^{r_1} \exp\left(-\delta_1'^2/2\right) \mathrm{d}\delta_1' \tag{2.38}$$

for the upper bound of probability of the correct estimate. Thus

$$P_{cU} = \varPhi(r_1) \tag{2.39}$$

where the function $\varPhi(\cdot)$ is tabled in detail e.g. in [3].

2.2. Lower Bound

Here we shall introduce two ways of calculating the lower bound of the correct estimate probability which we denote as ${}^1P_{cL}$, ${}^2P_{cL}$:

1) Let us create an auxiliary system of N hyperlayers $\widetilde{\mathcal{V}}_i$, $i = 1, \ldots, N$, defined by pairs of parallel hyperplanes $\widetilde{\mathcal{U}}_i$ which have distances $\tilde{r}_i$ from the origin. Directions of individual hyperplanes are chosen so as to be successively perpendicular to the individual axes of the coordinate system $\mathbb{C}'$. Then, equations (2.22) for these hyperplanes have the form

$$\widetilde{\mathcal{U}}_i\colon \delta_i' = \pm \tilde{r}_i; \quad i = 1, \ldots, N. \tag{2.40}$$

513

The distances $\tilde{r}_i$ are chosen as follows

$$\tilde{r}_1 := r_1 \tag{2.41}$$

$$\tilde{r}_i := \sqrt{(r_i^2 - r_1^2)}\,; \quad i = 2, \ldots, N \tag{2.42}$$

Put

$$\mathscr{W}_L = \bigcap \tilde{\mathscr{V}}_i \subset \mathscr{W}. \tag{2.43}$$

Then

$$^1P_{cL} = \int \ldots \int_{\mathscr{W}_L} p(\boldsymbol{\delta'})\,\mathrm{d}\boldsymbol{\delta'} = \prod \int_{-\tilde{r}_i}^{\tilde{r}_i} p(\delta'_i)\,\mathrm{d}\delta'_i. \tag{2.44}$$

We obtain a simple equation for the lower bound

$$^1P_{cL} = \Phi(r_1) \prod_{i=2}^{N} \Phi(\sqrt{(r_i^2 - r_1^2)}) \tag{2.45}$$

Condition (2.43) is very strict and can lead to an unnecessarily pessimistic lower bound, particularly if the values $\tilde{r}_2, \ldots, \tilde{r}_N$ are very near to r_1. This situation is geometrically illustrated in Fig. 2 for $N = 2$.
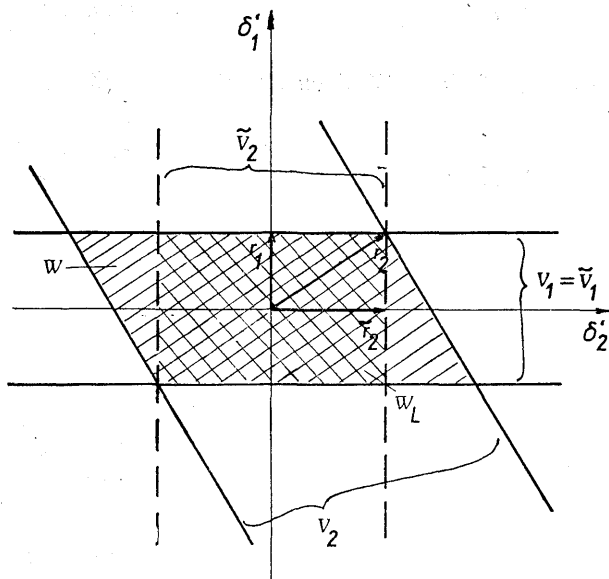
Fig. 2. Construction of domain $\mathscr{W}_L$. $\mathscr{W} = \mathscr{V}_1 \cap \mathscr{V}_2$; $\mathscr{W}_L = \tilde{\mathscr{V}}_1 \cap \tilde{\mathscr{V}}_2 \subset \mathscr{W}$.

2) Now we shall introduce another way of obtaining the lower bound $^2P_{cL}$. We shall take advantage of the directional invariability of distribution (2.30). Let us have volumes $\mathscr{W}$ and $\hat{\mathscr{W}} \subset \mathbb{R}_N$. Denote

$$\Delta\hat{\mathscr{W}} = \hat{\mathscr{W}} \setminus \mathscr{W} \tag{2.46}$$

$$\Delta\mathscr{W} = \mathscr{W} \setminus \hat{\mathscr{W}} \tag{2.47}$$

see Fig. 3. We can write

$$P_c = \int \ldots \int_{\mathscr{W}} p(\delta)\,\mathrm{d}\delta = \int \ldots \int_{\mathscr{W} \cap \hat{\mathscr{W}}} p(\delta)\,\mathrm{d}\delta + \int \ldots \int_{\varDelta\mathscr{W}} p(\delta)\,\mathrm{d}\delta \qquad (2.48)$$

Fig. 3. Sets $\varDelta\mathscr{W}$ and $\varDelta\hat{\mathscr{W}}$.

according to (2.25). If we select $\hat{\mathscr{W}}$ in such a manner that it holds

$$\int \ldots \int_{\varDelta\mathscr{W}} p(\delta)\,\mathrm{d}\delta \geqq \int \ldots \int_{\varDelta\hat{\mathscr{W}}} p(\delta)\,\mathrm{d}\delta \qquad (2.49)$$

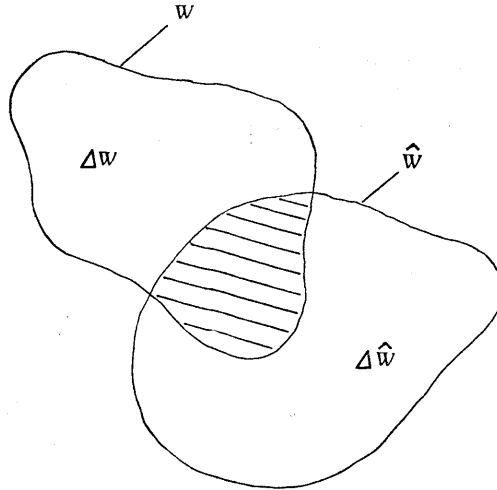then we obtain the lower bound $^2P_{cL}$ in the form

$$^2P_{cL} = \int \ldots \int_{\hat{\mathscr{W}}} p(\delta)\,\mathrm{d}\delta \leqq P_c . \qquad (2.50)$$

The value of p.d.f. (2.30) at any point $\delta \in \mathbb{R}_N$ depends only on the distance $\sqrt{(\sum \delta_i^2)}$ from the origin and does not depend on the direction. It follows from there that

$$\int \ldots \int_x p(\delta)\,\mathrm{d}\delta = \int \ldots \int_y p(\delta)\,\mathrm{d}\delta \qquad (2.51)$$

if the area y arose by a rotation of the area x around the origin or by folding x with respect to some radial axis. This is illustrated in Fig. 4 for $N = 2$. For simplicity, only halves of areas $\varDelta\mathscr{W}$, $\varDelta\hat{\mathscr{W}}$ are indicated in Fig. 4 and they are denoted schematically as $\frac{1}{2}\varDelta\mathscr{W}$, $\frac{1}{2}\varDelta\hat{\mathscr{W}}$. The second halves are symmetrical to them with respect to the origin. We can see that

$$\hat{\mathscr{W}} = \mathscr{V}_1 \cap \mathscr{V}_2 \qquad (2.52)$$

$$x = (\hat{\mathscr{V}}_2 \cap \mathscr{V}_1) \setminus \mathscr{W} = \varDelta\hat{\mathscr{W}} \qquad (2.53)$$

$$y \subset \varDelta\mathscr{W} = \mathscr{W} \setminus \hat{\mathscr{W}} \qquad (2.54)$$

515

in Fig. 4, where the hyperlayers $\hat{\mathscr{U}}_2$ are perpendicular to the axis δ_2 and they have distance r_2 from the origin. Then it holds (recorded symbolically for brevity)

$$^2P_{cL} := \int \ldots \int_{\hat{\mathscr{W}} \cap \mathscr{W}} + \int \ldots \int_{y} \leqq \int \ldots \int_{\hat{\mathscr{W}} \cap \mathscr{W}} + \int \ldots \int_{\Delta \mathscr{W}} = P_c \qquad (2.55)$$

Fig. 4. Construction of hyperlayer $\hat{\mathscr{V}}_2$ and relation between $\mathscr{W}$ and $\hat{\mathscr{W}}$.
$$\mathscr{W} = \mathscr{V}_1 \cap \mathscr{V}_2; \quad \hat{\mathscr{W}} = \mathscr{V}_1 \cap \hat{\mathscr{V}}_2.$$

If

$$r_3 \geqq \sqrt{(r_1^2 + r_2^2)} \qquad (2.56)$$

then it follows from (2.55), (2.44) and (2.45) for $N = 2$

$$^2P_{cL} = \Phi(r_1)\,\Phi(r_2) \qquad (2.57)$$

If (2.56) does not hold, then

$$^2P_{cL} = \Phi(r_1)\,\Phi(\sqrt{(r_3^2 - r_1^2)}). \qquad (2.58)$$

This approach can also be generalized, without difficulties, to N-dimensional space. Let the values $(r_2^*, r_3^*, \ldots, r_N^*)$ correspond to the smallest $N - 1$ values of elements of the set $\{r_2, r_3, \ldots, r_N, \tilde{r}_{N+1}, \ldots, \tilde{r}_{2N-1}\}$. Then, analogically as in (2.57), (2.58) we obtain the following equation for the lower bound

$$^2P_{cL} = \Phi(r_1)\prod_{i=2}^{N} \Phi(r_i^*) \qquad (2.59)$$

and it apparently holds

$$^1P_{cL} \leqq \, ^2P_{cL} \leqq P_c \leqq P_{cU}. \qquad (2.60)$$

2.3. Average Value and Dispersion of Correct Estimate

Let X be a nonrandom quantity. Then the average value of the correct estimate is

$$E_c = \mathsf{E}\{(1/a) \sum (X + \varDelta_i)/\sigma_{\varDelta_i}^2\} \qquad (2.61)$$

and after rearrangement

$$E_c = X + \mathsf{E}\{\sum \varDelta_i/\sigma_{\varDelta_i}^2\}/a = X \qquad (2.62)$$

since $\varDelta_i$ are zero-mean random variables. We can see that the correct estimate is unbiased.

The random variables $\varDelta_i$ are statistically independent and, therefore, the dispersion of the correct estimate is

$$\sigma_c^2 = \sigma^2\{(1/a) \sum \varDelta_i/\sigma_{\varDelta_i}^2\} \qquad (2.63)$$

and after rearrangement

$$\sigma_c^2 = 1/a . \qquad (2.64)$$

In a special case when all $\varDelta_i$ have the same variances $\sigma_{\varDelta_i}^2 = \sigma_\varDelta^2$; $i = 1, 2, ..., N$, then

$$\sigma_c^2 = \sigma_\varDelta^2/N . \qquad (2.65)$$

If X is a random variable with the uniform p.d.f., then equations (2.61) to (2.65) also hold, but the moments occurring there should be considered as conditional moments for the given X.

3. QUALITY OF OPTIMUM ESTIMATE WITH A PRIORI INFORMATION

The optimum estimation rule is given by (1.13) and (1.14) in this case. Similarly to the case without a priori information, this estimation rule can be expressed in the form: Calculate a pair of values $\tilde{X}$, h

$$\tilde{X} = (\vartheta/\sigma_\vartheta^2 + \sum d_i/\sigma_{\varDelta_i}^2)/A = \tilde{X}(\boldsymbol{d}) \qquad (3.1)$$

$$h = \sum [a(\tilde{X} - d_i)^2 + (\vartheta - d_i)^2/\sigma_\vartheta^2]/\sigma_{\varDelta_i}^2 = h(\boldsymbol{d}) \qquad (3.2)$$

for all vectors $\boldsymbol{d} \in \mathcal{D}$, where

$$A = a + 1/\sigma_\vartheta^2 . \qquad (3.3)$$

Only the value $\tilde{X} \in \{X(\boldsymbol{d})\}$ is the optimum estimate $\hat{X}$, to which the least value $h(\boldsymbol{d})$; $\boldsymbol{d} \in \mathcal{D}$, belongs.

Further we shall use the following notation

$$\tau = (\sum m_i \lambda_i/\sigma_{\varDelta_i}^2)/A \qquad (3.4)$$

$$\mu_j = (\tau - m_j \lambda_j)/\sigma_{\varDelta_j}^2 ; \quad j = 1, ..., N . \qquad (3.5)$$

If we substitute

$$\vartheta = X - \varDelta_\vartheta \tag{3.6}$$

$$d_i = D_i + m_i\lambda_i = X + \varDelta_i + m_i\lambda_i \tag{3.7}$$

into (3.1) and (3.2) for $m = 0$, we obtain a pair of quantities $\tilde{X}_T$, h_T for the trunk of estimates. If $m \neq 0$, we obtain a pair $\tilde{X}_B$, h_B corresponding to branches. The equation for the quantity h_B can be arranged into the form

$$h_B = h_T + Y + Z ; \quad m \neq 0 \tag{3.8}$$

where

$$Y = \sum_{i=1}^{N+1} \alpha_i \delta_i \tag{3.9}$$

$$\alpha_i = 2[(a/A\sigma_{\varDelta_i}) \sum_j \mu_j - a\mu_i\sigma_{\varDelta_i} + m_i\lambda_i/\sigma_\vartheta^2\sigma_{\varDelta_i}] = \text{const.,} \quad i = 1,\dots,N \tag{3.10}$$

$$\alpha_{N+1} = 2[\overset{\centerdot}{-}a\sigma_\vartheta(1 - a/A) \sum \mu_i + (\sum m_i\lambda_i/\sigma_{\varDelta_i}^2)/\sigma_\vartheta] = \text{const.} \tag{3.11}$$

$$\delta_i = \varDelta_i/\sigma_{\varDelta_i} ; \quad i = 1,\dots,N \tag{3.12}$$

$$\delta_{N+1} = \varDelta_\vartheta/\sigma_\vartheta \tag{3.13}$$

$$Z = (a/A)[(a/A) - 2](\sum m_i\lambda_i/\sigma_{\varDelta_i}^2)^2 + A \sum (m_i\lambda_i/\sigma_{\varDelta_i})^2 = \text{const.} \tag{3.14}$$

All quantities δ_i, $i = 1,\dots,(N+1)$, are independent Gaussian random variables $\mathcal{N}(0, 1)$.

Let us compare (3.8) and (3.9) with the analogical equations (2.7) and (2.8) for the case where no a priori information exists. Comparing them, we can see the formal conformity of both cases. The only difference is an increased extent of summation in (3.9). The manner of determining the correct-estimate probability and its bounds in the case where a priori information exists is the same as described and discussed in Section 2. But we must remember that we deal with the vectors in the space $\mathbb{R}_{N+1}$ instead of $\mathbb{R}_N$ and the constants α_i, Z are given by (3.10), (3.11) and (3.14) instead of (2.9), (2.10).

3.1. Average Value and Dispersion

We shall develop these measures separately for the case where X is a nonrandom quantity whilst an estimate uses an additional observation ϑ and separately for the case where X is a random quantity.

In the first case we shall seek moments for the quantity

$$\hat{X} = [(X - \varDelta_\vartheta)/\sigma_\vartheta^2 + \sum (X + \varDelta_i)/\sigma_{\varDelta_i}^2]/A . \tag{3.15}$$

Then

$$E_c = X(1/\sigma_\vartheta^2 + a)/A = X \tag{3.16}$$

$$\sigma_c^2 = (\sigma_\vartheta^2/\sigma_\vartheta^4 + \sum \sigma_{\varDelta_i}^2/\sigma_{\varDelta_i}^4)/A^2 = 1/A . \tag{3.17}$$

518

In the second case we shall seek conditional moments (for given X) for the quantity

$$\hat{X} = [m_X/\sigma_X^2 + \sum(X + \Delta_i)/\sigma_{\Delta_i}^2]/A' \tag{3.18}$$

where

$$A' = a + 1/\sigma_X^2 . \tag{3.19}$$

Then

$$E_c = (m_X/\sigma_X^2 + aX)/A' \tag{3.20}$$

$$\sigma_c^2 = \sum(\sigma_{\Delta_i}/\sigma_{\Delta_i}^2)^2/A'^2 = a/A'^2 . \tag{3.21}$$

The estimate is biased now though it has the lesser variance (for $A = A'$) than in the first case.

4. EXAMPLE

Here we shall introduce a simple example of calculating the quality of the optimum estimation rule without a priori information for the case $N = 2$, $\lambda_1/\lambda_2 = \frac{4}{3}$.

First we shall determine the set $\mathcal{M}$ of vectors $\boldsymbol{m} = (m_1, m_2)$. Decompose $\mathcal{M}$ at two subsets $\mathcal{M}^+$, $\mathcal{M}^-$; $\mathcal{M}^+ \cup \mathcal{M}^- = \mathcal{M}$. The set $\mathcal{M}^+$ contains only vectors $\boldsymbol{m}$ with non-negative components m_1, m_2. Evidently, $\mathcal{M}^-$ contains vectors $\boldsymbol{m}$ which are complementary to elements of $\mathcal{M}^+$. We can, therefore confine ourselves to e.g. $\mathcal{M}^+$. In our case (see Fig. 5), considering the possible errors Δ_i, the set $\mathcal{M}^+$ has elements $\mathcal{M}^+ = \{(0, 1), (1, 0), (1, 1), (1, 2), (2, 1), (2, 2)\}$.
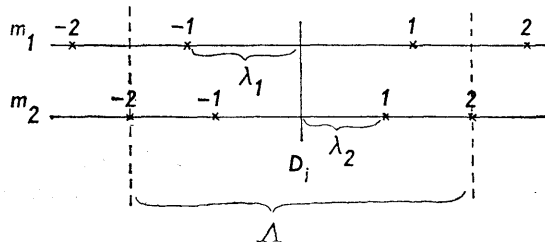

Fig. 5. Determining of set $\mathcal{M}$.

For each $\boldsymbol{m} \in \mathcal{M}^+$ we calculate $\alpha_1 = \alpha_1(\boldsymbol{m})$ and $\alpha_2 = \alpha_2(\boldsymbol{m})$ according to (2.9) and from there we calculate $r(\boldsymbol{m})$ according to (2.23). The obtained quantities $r(\boldsymbol{m})$ are then arranged into a sequence in accordance with their increasing values. Thus, we get a sequence $\{r_j\}$ and we calculate the corresponding sequences $\{\tilde{r}_j\}$ and $\{r_j^*\}$ according to (2.42) and to the text following (2.58).

Table 1 displays results of calculations constant phase-variances $\sigma^2 = \sigma_{\Delta_i}^2/\lambda_i^2 = $ $= $ const., $i = 1, 2$. We can see from Table 1 that $r_1^* = r_1$ and $r_2^* = \tilde{r}_3$. Now, we can calculate the bounds of probability of correct estimate P_{cU} and P_{cL} according to (2.39) and (2.59). (It also holds $^2P_{cL} = {}^1P_{cL}$ in our case since $r_2^* = \tilde{r}_2$.) The calculated values of P_{cU} and P_{cL} are displayed in Table 2 for several values of σ. Also,

values of P_{sim} are indicated there for comparison, where P_{sim} is the probability of the correct estimate obtained by a simulation method. The extent of the simulation was 10 000 trials. Table 2 further introduces the normalized standard deviation of the correct estimate $\sigma_n = \sigma_c / \mathsf{E}\{\lambda_i\}$. The average value $\mathsf{E}\{\hat{X} - X\}$ is zero.

Table 1. Calculation of sequences $\{r_j\}$ and $\{\tilde{r}_j\}$.

m_1	1	2	2	1	0	1
m_2	1	1	2	0	1	2
j	1	2	3	4	5	6
$\alpha_1\sigma$	$-0{\cdot}24$	$0{\cdot}48$	$-0{\cdot}48$	$0{\cdot}72$	$-0{\cdot}96$	$-1{\cdot}20$
$\alpha_2\sigma$	$0{\cdot}32$	$-0{\cdot}64$	$0{\cdot}64$	$-0{\cdot}96$	$1{\cdot}28$	$1{\cdot}60$
$r_j\sigma$	$0{\cdot}1$	$0{\cdot}2$	$0{\cdot}2$	$0{\cdot}3$	$0{\cdot}4$	$0{\cdot}5$
$\tilde{r}_j\sigma$	$0{\cdot}1$	$0{\cdot}173$	$0{\cdot}173$	$0{\cdot}283$	$0{\cdot}387$	$0{\cdot}49$

Table 2. Calculation of the estimate quality for several values σ.

j	$r_j^*\sigma$	$\sigma = 0{\cdot}03$		$0{\cdot}06$		$0{\cdot}10$	
		r_j^*	$\phi(r_j^*)$	r_j^*	$\phi(r_j^*)$	r_j^*	$\phi(r_j^*)$
1	$0{\cdot}1$	$3{\cdot}33$	$0{\cdot}999$	$1{\cdot}67$	$0{\cdot}905$	$1{\cdot}0$	$0{\cdot}68$
2	$0{\cdot}173$	$5{\cdot}77$	$1{\cdot}0$	$2{\cdot}89$	$0{\cdot}996$	$1{\cdot}73$	$0{\cdot}92$
	P_{cU}		$0{\cdot}999$		$0{\cdot}905$		$0{\cdot}68$
	P_{sim}		$0{\cdot}999$		$0{\cdot}903$		$0{\cdot}67$
	P_{cL}		$0{\cdot}999$		$0{\cdot}901$		$0{\cdot}62$
	σ_n		$0{\cdot}025$		$0{\cdot}050$		$0{\cdot}071$

It follows from the theory and it is also confirmed by calculations that the greater are the values r_j^*; $j = 2, \ldots, N$, the closer the upper P_{cU} and lower P_{cL} bounds are to each other.

CONCLUSIONS

We have developed a method of calculating the quality of an estimation. A seemingly complicated development leads to simple final equations for the bounds of probability of the correct estimate.

ACKNOWLEDGEMENT

REFERENCES

[1] I. Vrana: Optimum statistical estimates in conditions of ambiguity. IEEE Trans. Inform. Theory *IT-38* (in print).
[2] J. W. Dettman: Mathematical Methods in Physics and Technic. McGraw-Hill, New York 1962.
[3] Tables of Error Function and Its First Twenty Derivatives. Harvard University Press, Cambridge 1952.

Doc. Ing. Ivan Vrana, DrSc., katedra informatiky Provozně ekonomické fakulty VŠZ (Department of Informatics, Faculty of Operational — Economy, University of Agriculture), 160 21 Prague 6–Suchdol. Czechoslovakia.