OPTIMAL STRATEGY IN CORRECTION

Corresponding Member of the USSR Academy of Sciences D. E. OKHOTSIMSKII,

1967

SovietRxiv

Abstract

Full Text

UDC 62–50:531.55

CYBERNETICS AND CONTROL THEORY

Corresponding Member of the USSR Academy of Sciences D. E. OKHOTSIMSKII,
V. A. RYASIN, N. N. CHENTSOV

OPTIMAL STRATEGY IN CORRECTION

This note is devoted to the problem of constructing an optimal strategy for correcting a vehicle moving near a nominal trajectory. In contrast to works ([1–3]), where strategies are considered with correction times that do not depend on the results of trajectory measurements, here flexible strategies are considered, in which the correction times and the magnitudes of the correcting impulses are assigned during flight according to the results of the trajectory measurements that have been carried out. Instead of the criterion of minimizing the mathematical expectation of fuel expenditure, a direct criterion is adopted: maximizing the probability of hitting a prescribed region of the space of corrected parameters for a given limited fuel reserve.

1. The basis of the consideration is the following linearized model. The motion of the vehicle in a neighborhood of the nominal trajectory is described by the system of differential equations:

$$\dot{\mathbf{h}}(t) = R(t)\mathbf{h}(t) + R_1(t)\mathbf{u}(t),$$

where $\mathbf{h}(t)$ is the six-dimensional vector of phase deviations of the actual trajectory from the nominal one, $\mathbf{u}(t)$ is a control consisting of a finite number of control impulses. A control impulse changes the velocity of the vehicle, leaving its coordinates unchanged. It is assumed that execution errors are absent. The one-dimensional space of quantities δ, which are linear combinations of the components of the vector $\mathbf{h}(T)$, is called the space of corrected parameters or the space of misses. In this space an interval of admissible misses $A = (-\delta_0, \delta_0)$ is chosen. It is assumed that the correction has been successful if at the time T the event

$$\{\delta \in A\}. \tag{1}$$

occurs.

At fixed instants of time $t_1, \dots, t_N$, trajectory measurements are carried out. The deviations α_i of certain quantities from their nominal values are measured; these deviations depend linearly on the insertion errors $\mathbf{h}(0)$ and on the control over the time interval preceding the measurement. The measurements are made with errors Δ_i. The result of a measurement is written in the form

$$x_i = b_i(\mathbf{u}[0, t_i]) + l_i \cdot \mathbf{h}(0) + \Delta_i = \alpha_i + \Delta_i,$$

where b_i is the term characterizing the influence on α_i of the control on the interval $[0, t_i]$; l_i is a vector associated with the time t_i. The space X of sets $x = (x_1, \dots, x_N)$ is called the sample space. It is assumed that on the space of insertion errors and measurement errors (the space Ω of elementary outcomes) a Gaussian probability is specified.

A strategy is a rule that makes it possible, for each sample x of observations, to assign the correction times and the correcting impulses. The decision on each correction is made on the basis of the actual trajectory measurements, the number and accuracy of future trajectory measurements, the available fuel reserve, and the number of corrections whose execution is assumed in the future. In the note, strategies with one and two corrections are considered. A strategy with one correction is defined by a partition of the spac-

of the equality of samples on N cylindrical sets $B_1, \dots, B_N$ and by specifying on each B_i a time function $t_i \leq \tau(\mathbf{x}_i) < t_{i+1}$ and a control function $|\mathbf{q}(\mathbf{x}_i)| \leq W_0$, where W_0 is the fuel reserve for correction, $\mathbf{x}_i = (x_1, \dots, x_i)$. A strategy in two-phase correction is determined by a partition of the sample space into cylindrical sets B_{ij}, $i = 1, 2, \dots, N-1$, $i < j \leq N$, on which time functions and control functions are specified for the first correction $t_i \leq \tau_1(\mathbf{x}_i) < t_{i+1}$, $\mathbf{q}_1(\mathbf{x}_i)$, and for the second correction $t_j \leq \tau_2(\mathbf{x}_j, i) < t_{j+1}$, $\mathbf{q}_2(\mathbf{x}_j, i)$, in such a way that $|\mathbf{q}_1| + |\mathbf{q}_2| \leq W_0$.

To each strategy ν there corresponds a definite probability $P(\nu)$ of event (1). A strategy ν^* is optimal in the given class if, for any other strategy ν from this class, $P(\nu) \leq P(\nu^*)$. The problem consists in constructing optimal strategies in the classes E_1 and E_2 of strategies with one and with two corrections.

2. From the classes E_1 and E_2 we single out the so-called complete subclasses Π_1 and Π_2 of strategies, which have an advantage in the manner of specification and contain an optimal strategy. Consider the influence function $\theta(t)$, showing what maximum shift in the miss space can be obtained by a correction at time t, having a unit fuel reserve. We shall denote by $\vec{\theta}(t)$ the direction of the impulse leading to the maximum shift. Let $\theta(t)$ be continuous. By θ_i we denote the maximum of $\theta(t)$ on $[t_i, t_{i+1})$. We choose times τ_i satisfying the following conditions: $t_i \leq \tau_i < t_{i+1}$, $\theta(\tau_i) = \theta_i$, $\theta(t) < \theta_i$ for $t > \tau_i$. There are no more than one such times on each interval $[t_i, t_{i+1})$, and if the influence function is not monotone, then on some intervals they may be absent.

By definition, the class Π_m, $m = 1, 2$, consists of those and only those strategies from E_m which allow correction only at the times τ_i in the direction $\vec{\theta}(\tau_i)$.

Lemma. For every strategy $\nu \in E_m$ there exists a strategy $\rho \in \Pi_m$ such that $P(\nu) = P(\rho)$, and requiring a smaller expenditure of fuel.

On the basis of the lemma, Π_m forms a complete class, to which we shall restrict ourselves in the subsequent reasoning.

3. The partition of the sample space together with the totality of functions of the time and magnitude of the impulse represents a way of specifying a strategy as a function on the space of elementary outcomes. The classes of strategies under consideration on Ω are conveniently specified with the aid of another sample space $Z = \{z_1 \dots z_N\} = \{z\}$, whose coordinates are the differences of miss forecasts under successive measurements in the absence of control. By the forecast of the miss $z(i)$ at time t_i under conditions of absent control is meant the mathematical expectation of the miss for given $x_1, \dots, x_i$, if on $[0, t_i]$ the control is equal to zero. If, however, on $[0, t_i]$ the control is not equal to zero, then in the definition of the forecast $x_1, \dots, x_i$ should be replaced by $y_1, \dots, y_i$, where $y_i = l_j \cdot h(0) + \Delta_j = x_j - b_j(\mathbf{u}[0, t_j])$. It is proved that the class of strategies under consideration can be specified with the aid of Z.

4. Consider a one-time correction. To simplify the notation, we shall assume that the times τ_i possible for carrying out corrections are present on every interval $[t_i, t_{i+1})$. The generalization to the case of a smaller number of possible correction times is obvious.

For a strategy with one correction we write out the expression for the probability of success of the correction:

$$P(\nu) = \sum_k \int_{B_k} \left[\int_A \frac{1}{\sqrt{2\pi}\,\sigma_k} \exp\left\{ -\frac{1}{2\sigma_k^2} \left(\delta - z(k) - q(z_k)\theta_k\right)^2 \right\} d\delta \right] d\mathcal{F}^0(z_k),$$

where $\sigma_k^2 = \sigma_0^2 - Mz_1^2 - \dots - Mz_k^2$, and σ_0^2 is the a priori variance of the miss. In the square brackets stands the probability of success $\varphi(z(k), q(z_k))$ for given $z_k, q(z_k)$.

As the control function at the instant τ_i we consider

$$Q(z_k) = \begin{cases} -z(k)/\theta_k, & \text{if } |z(k)| \le W_0\theta_k, \\ -\operatorname{sign} z(k) \cdot W_0, & \text{if } |z(k)| > W_0\theta_k. \end{cases} \tag{2}$$

This function maximizes the value of the conditional probability of successful correction for each sample z_k. Hence, for any partition $\{B_k\}$ of the space Z, as the control functions one should always take $Q(z_k)$. The optimal partition

Fig. 1. One-time correction. The region D_k for each instant τ_k is the part of the straight line $W = W_0$ for $|a| > a_k$.

Figure 1: Fig. 1. One-time correction. The region D_k for each instant τ_k is the part of the straight line $W = W_0$ for $|a| > a_k$.

$\{B_{k\ \text{opt}}\}$ can be found by the method of dynamic programming. Denote $\psi(z_k) = \varphi(z(k), Q(z_k))$, $r(z_N) = \psi(z_N)$,

$$r(z_k) = \max\left\{\psi(z_k), \int r(z_{k+1})\, d\mathcal{P}(z_{k+1})\right\}$$

and introduce the sets

$$C_k = \left\{z : \psi(z_k) > \int r(z_{k+1})\, d\mathcal{P}(z_{k+1})\right\}.$$

Then $\{B_{k\ \text{opt}}\}$ is written in the form $C_1,\ \overline{C}_1 \cap C_2, \dots, \dots, (C_1 \cup \dots \cup C_{N-1})$. The optimal control function (2) at the instant τ_k, for a fixed fuel reserve W_0, depends only on the forecast and not on the entire system of measurements $z_1, \dots, z_k$. Hence $\psi(z_k)$ depends on the forecast. The formulas for $r(z_N)$ and $r(z_k)$ make it possible to conclude by induction that $r(z_k)$ also depends only on the forecast. Denote $\psi(z_k)$ and $r(z_k)$ by $\lambda_k(a)$ and $\mu_k(a)$, where $a = z(k)$, and introduce on the numerical axis the sets:

$$D_k = \left\{a : \lambda_k(a) > \int \mu_{k+1}(a + z_{k+1})\, d\mathcal{P}(z_{k+1})\right\}.$$

Then, as follows from the expression for C_k, the set C_k consists of those samples from Z for which the sum of the first k elements is a number from D_k: $C_k = \{z : z(k) \in D_k\}$. The decision-making process is as follows: 1) correction is assigned at the instant τ_k, when the forecast first enters the region D_k; 2) the magnitude of the correcting impulse is assigned according to formula (2). An illustrative graph is given in Fig. 1.

Fig. 1. One-time correction. The region D_k for each instant τ_k is the part of the straight line $W = W_0$ for $|a| > a_k$.

Consider the case of almost exact measurements, when $\sigma_i^2 - Mz_1^2 \ll \delta_0^2$. This condition means that the random displacement of the forecast at the instant τ_i, $i > 1$, from the forecast at the instant τ_1 is small on the scale σ_0, although it may be large on the scale of the interval A of admissible misses. The function $\lambda_k(a)$, under conditions of almost exact measurements, has the form of an almost step function. The region of the smeared jump is in the neighborhood of the point $a = W_0\theta_k$ and has magnitude of order $3\sigma_k$, small on the scale $3\sigma_0$. The regions D_k and the functions $\mu_k(a)$ are easily determined:

$$D_k \approx \{a : a > W_0\theta_{k+1}\}$$

with accuracy $3\sigma_k$,

$$\mu_k(a) = \begin{cases} \Phi(\delta_0/\sigma_k) - \Phi(-\delta_0/\sigma_k), & \text{if } W_0\theta_{i+1} \le |a| < W_0\theta_i, \\ & \quad i = k+1, \dots, N, \\ \lambda_k(a), & \text{if } |a| > W_0\theta_{k+1}, \end{cases}$$

where $\theta_{N+1} = 0$, $\Phi(x)$ is the distribution function of the standard normal law. The optimal probability of success is computed by the formula

$$P(\nu^*) = \int \mu_1(a)\, d\mathcal{P}(a)$$

and is approximately equal to

$$\sum_1^N \left[\Phi(\delta_0/\sigma_i) - \Phi(-\delta_0/\sigma_i)\right] \left[\Phi\left(W_0\theta_i/\sqrt{Mz_1^2}\right) - \Phi\left(-W_0\theta_{i+1}/\sqrt{Mz_1^2}\right)\right].$$

5. Consider two-time correction. The question of constructing an optimal strategy reduces to the question of constructing an optimal strategy in the subclass U_k of strategies with two corrections, for which the instant of the first correction is equal to τ_k with probability 1.

Theorem. *In the subclass U_k of strategies with two corrections and with a fixed time of the first correction, there exists an optimal one.*

The proof of the theorem is connected with approximation in the space of strategies.

It can also be shown that among the optimal strategies there is a strategy in which the control functions for the first impulse $Q_1(a)$ depend only on the miss forecast $a = z(k)$. For such a strategy, on the line one introduces sets depending on the fuel reserve, $D_l''(W)$ and $D_k'(W)$, which, as in the case of a one-time correction, determine the optimal partition $\{B_{k\,\mathrm{opt}}\}$ and $\{B_{kl\,\mathrm{opt}}\}$ of the space Z. The method of behavior under a two-phase correction is as follows. The coordinates z are summed. As soon as the event $\{z(k) \in D_k'(W_0)\}$ occurs, the first correction is carried out by an impulse of magnitude $Q_1(z(k))$. Subsequent measurements are made, and the second correction is assigned as soon as the event $\{a' \in D_l''(W_0 - |Q_1(z(k))|)\}$ occurs, where

$$a' = z(k) + Q_1(z(k))\theta_k + z_{k+1} + \dots + z_l.$$

The control function for the second impulse is obtained if in (2) the value W_0 is replaced by $W_0 - |Q_1(z(k))|$. Calculations show that, if the measurements are not nearly exact, then the optimal control function for the first impulse at the time τ_k is, in absolute value, smaller than function (2), i.e., undercorrection takes place. With nearly exact measurements, undercorrection disappears. In this case many methods close to the optimal one can be proposed. The simplest consists in the following: after the first measurement an impulse $Q_1 = z(1)/\theta_1$ is imparted. The second correction is carried out at the time τ_N: $Q_2 = z(N)/\theta_N$. The probability of success under such a strategy is approximately equal to

$$[\Phi(\delta_0/\sigma_N) - \Phi(-\delta_0/\sigma_N)][\Phi(W_0\theta_1/\sqrt{Mz_1^2}) - \Phi(-W_0\theta_1/\sqrt{Mz_1^2})].$$

The error in the probability of success in comparison with the optimal one tends to zero as the accuracy of the measurements increases.

Received
9 II 1967

CITED LITERATURE

[1] J. V. Breakwell, F. Tung, R. R. Smith, AIAA J., **3**, No. 5, 807 (1965).
[2] F. Tung, IEEE Trans. on Automatic Control, AC-10, No. 3, 328 (1965).
[3] V. A. Yaroshevsky, T. V. Parysheva, *Cosmic Research*, **3**, issue 6, 826 (1965); **4**, issue 1, 826 (1966).
[4] V. A. Ryasin, *Theory of Probability and Its Applications*, **11**, issue 4, 708 (1966).

Note: Figure translations are in progress. See original paper for figures.