KYBERNETIKA ČÍSLO 3, ROČNÍK 3/1967

The Optimal Control of Linear System with Random Stationary Perturbations

OTAKAR ŠEFL

This paper deals with the optimal control of a linear system with an additive random perturbation. The control function is to be estimated on the basis of the observation of a signal, which is statisticaly dependend on the random perturbation and thus the control-function should have the prescribed dispersion. It is proved that under certain conditions the control-function can be obtained as the solution of a certain functional equation.

1.

By a random system we call a plant, which can be mathematically described as a random transformation of the input- on to output-signals. It means that for a given input the output is a random variable or a random process. The properties of such systems are usually specified by a class of conditional probabilities, which described the probability distribution of the output under the condition of the given input. The mathematical description of a random system of this type is very universal and complete (we cannot obtain more information on such systems than those which are involved in the class of conditional probabilities), but it has some disadvantages. Excluding the case of normal distribution it is very difficult or even impossible to obtain more detailed solution in analytical form and also the physical determination of the conditional probabilities is a difficult problem.

From these reasons it seems useful to us to use in certain cases for the description of a random system a combination of classical and statistical methods. In this paper we shall study a system of this type.

Fig. 1 shows the scheme of this system.

The output of the controlled plant is denoted by x and it equals $x = y(u) - s$, where $y(u)$ is a linear transformation of the input signal and s is a random perturbation. The input u is the output of the regulator. The regulator has only one input denoted by w. This signal w depends statistically on s resp. on x. (We shall see later that for the solution of our problem the knowledge of the conditional mean value of s under condition w is sufficient). We shall try to estimate the control-function u, which minimizes the dispersion of x and which has the second moment less than a given number.

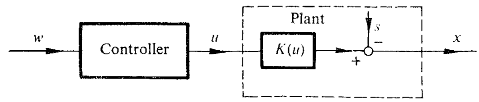

Fig. 1.

2.

Now we shall give a mathematical formulation of our problem. We shall study a discrete model of our problem. It means that all signals in our system are described by sequences of random variables or numbers. In what follows we denote by capital letters $X, S, U, W, \ldots$ random variables and by corresponding small letters $x, s, u, w, \ldots$ their values. $\mathscr{X} = \{X_n\}$, $\mathscr{U} = \{V_n\}, \ldots$ denote discrete stochastic processes and the corresponding $x = \{x_n\}$, $u = \{u_n), \ldots$ are their realisations. If $\mathscr{X} = \{X_n\}$ is a stationary stochastic process for which the mean value $E\{X_n\} = 0$ and the dispersion $E\{X_n^2\}$ is finite, then $\hat{X}(\nu)$ denotes its spectral representation, i.e.

$$X_n = \int_{-\pi}^{\pi} e^{i\nu n} \, d\hat{X}(\nu) . \tag{2.1}$$

The spectral function corresponding to the process $\mathscr{X}$ will be denoted by $\gamma_{\mathscr{X}}(\nu)$.

If $\mathscr{X}$ and $\mathscr{Y}$ is a stationary couple of stochastic processes. then we define

$$(\mathscr{X}, \mathscr{Y}) = E\{X_n Y_n\} \tag{2.2}$$

and

$$(\mathscr{X}, \mathscr{X}) = \|\mathscr{X}\|^2 . \tag{2.3}$$

The random perturbation and the measured signal are assumed to be realisations of a stationary couple of random processes, which we denote by $\mathscr{S} = \{S_n\}$ and $\mathscr{W} = \{W_n\}$ and which fulfil the following conditions

$$\begin{gathered} E\{S_n\} = E\{W_n\} = 0 , \\ E\{S_n^2\} < \infty , \quad E\{W_n^2\} < \infty , \\ E\{S_n^2 \mid W_n, W_{n-1}, \ldots\} < \infty . \end{gathered} \tag{2.4}$$

Let $\mathscr{F} = \{f\}$ be the set of all functionals on the set of all sequences for which $f(W_n, W_{n-1}, \ldots)$ is a random variable;

$$E\{f^2(W_n, W_{n-1}, \ldots\} \leqq c < \infty . \tag{2.5}$$

Then we define the control function as a random process $\mathcal{U}_f = \{U_n\}$, where $U_n = f(W_n, W_{n-1}, \ldots)$ for every integer n and $f \in \mathcal{F}$. The class of all $\{\mathcal{U}_f\}_{f\in\mathcal{F}}$ we call class of all admissible control functions (a.c.f.).

Denote now by $\mathcal{Y}(\mathcal{U}_f)$ the process $\{Y_n\}$ on the „output" of the linear part of the plant and let

$$Y_n = \int_{-\pi}^{\pi} e^{i\nu n} K(e^{i\nu})\, dU(\nu) \tag{2.6}$$

where $K(e^{i\nu})$ is a rational function

$$K(e^{i\nu}) = \frac{\sum_{k=0}^{q} a_k e^{i\nu k}}{\sum_{k=0}^{r} b_k e^{i\nu k}}\,. \tag{2.7}$$

Let for this function $r > q$ and $\left|\sum_{k=0}^{r} b_k z^k\right| > 0$ whenever $|z| \geqq 1$. (z denotes here a complex number.)

The output of the plant is a realisation of a stochastic process

$$\mathcal{X}(\mathcal{U}_f) = \mathcal{Y}(\mathcal{U}_f) - \mathcal{S}\,. \tag{2.8}$$

(The difference of two processes $\mathcal{Y}$ and $\mathcal{S}$ is defined as the process $\{Y_n - S_n\}$). We shall try to find a control-function $f^* \in \mathcal{F}$ such that

$$\|\mathcal{X}(\mathcal{U}_{f^*})\|^2 \leqq \|\mathcal{X}(\mathcal{U}_f)\|^2 \quad \text{for every} \quad f \in \mathcal{F}\,. \tag{2.9}$$

This control function we call the optimal control-function (o.c.f.).

3.

Now we give a sufficient condition for an a.c.f. to be optimal. To this purpose we introduce the following notation: $\mathcal{V} = \{V_n\}$ is a stochastic process, where

$$V_n = E\{S_n \mid W_n, W_{n-1}, \ldots\} \tag{3.1}$$

and $\mathcal{Y}^{-1}(\mathcal{U})$ denotes the stochastic process, whose members are defined by

$$\int_{-\pi}^{\pi} e^{i\nu n}\, \overline{K^{-1}(e^{i\nu})}\, d\, \hat{U}(\nu) \tag{3.2}$$

where $K^{-1}(e^{i\nu})$ is the inverse function to the function defined in (2.7). In order the that the expression (3.2) is senseful we have to assume that

$$\int_{-\pi}^{\pi} |K^{-1}(e^{i\nu})|^2\, d\, \gamma_{\mathcal{U}}(\nu) < \infty\,. \tag{3.3}$$

Now we can prove the following theorem.

Theorem 3.1. *Let there exist such a a.c.f. f and such a real number $\lambda > 0$ that*

$$(\mathscr{Y}(\mathscr{U}_f) + \lambda\mathscr{Y}^{-1}(\mathscr{U}_f) = \mathscr{V}\,, \tag{3.4}$$

$$\|\mathscr{U}_f\|^2 = c\,, \tag{3.5}$$

$$\lambda > 0\,. \tag{3.6}$$

Then f is a optimal control function.

Proof. Let f^* be a a.c.f. which fulfils all conditions of the theorem 3.1 and let g be any other a.c.f. Write g in the form $g = f^* + h$. From the definition of $\mathscr{U}_f$ it follows that $\mathscr{U}_{f^*+h} = \mathscr{U}_{f^*} + \mathscr{U}_h$ and from the definition of $\mathscr{Y}$ (see 2.6) it is evident that

$$\mathscr{Y}(\mathscr{U}_{f^*+h}) = \mathscr{Y}(\mathscr{U}_{f^*}) + \mathscr{Y}(\mathscr{U}_h)\,. \tag{3.7}$$

According to the definition of $\mathscr{U}_f$ the random variable U_n depends for every n on $W_n, W_{n-1}, \ldots$ only. From these properties of $\mathscr{U}_f$ and from the definition of $\mathscr{V}$ it follows that

$$(S, \mathscr{Y}(\mathscr{U}_f)) = (\mathscr{V}, \mathscr{Y}(\mathscr{U}_f)) \tag{3.8}$$

for every $f \in \mathscr{F}$.

Because the function $g = f^* + h$ belongs to the class $\mathscr{F}$ it must hold $\|\mathscr{U}_{f^*+h}\|^2 \leqq \leqq c$. According to the assumption (3.5) it follows that

$$\|\mathscr{U}_{f^*+h}\|^2 \leqq \|\mathscr{U}_f\|^2\,. \tag{3.9}$$

It is clear that this relation holds only if

$$(\mathscr{U}_{f^*}, \mathscr{U}_h) \leqq 0\,. \tag{3.10}$$

Now let us make the difference

$$\|\mathscr{X}(\mathscr{U}_{f^*+h})\|^2 - \|\mathscr{X}(\mathscr{U}_{f^*})\|^2\,. \tag{3.11}$$

Substituting in this expression for $\mathscr{X}(\mathscr{U}_{f^*+h})$ and using (3.7) and (3.8) we obtain

$$\begin{aligned} &\|\mathscr{X}(\mathscr{U}_{f^*+h})\|^2 - \|\mathscr{X}(\mathscr{U}_{f^*})\|^2 = \\ &= \|\mathscr{Y}(\mathscr{U}_{f^*+h}) - \mathscr{S}\|^2 - \|\mathscr{Y}(\mathscr{U}_{f^*}) - \mathscr{S}\|^2 = \\ &= \|\mathscr{Y}(\mathscr{U}_{f^*}) + \mathscr{Y}(\mathscr{U}_h) - \mathscr{S}\|^2 - \|\mathscr{Y}(\mathscr{U}_{f^*}) - \mathscr{S}\|^2 = \\ &= 2(\mathscr{Y}(\mathscr{U}_{f^*}) - \mathscr{S}, \mathscr{Y}(\mathscr{U}_h)) + \|\mathscr{Y}(\mathscr{U}_h)\|^2 = \\ &= 2(\mathscr{Y}(\mathscr{U}_f) - \mathscr{V}, \mathscr{Y}(\mathscr{U}_h)) + \|\mathscr{Y}(\mathscr{U}_h)\|^2\,. \end{aligned} \tag{3.12}$$

According to the condition (3.4) the expression (3.12) equals

$$-2\lambda(\mathscr{Y}^{-1}(\mathscr{U}_{f^*})\,\mathscr{Y}(\mathscr{U}_{f^*})) + \|\mathscr{Y}(\mathscr{U}_h)\|^2\,. \tag{3.13}$$

From the definition of $\mathcal{Y}(\mathcal{U}_{f*})$ and $\mathcal{Y}^{-1}(\mathcal{U}_{f*})$ it is easy to see that $(\mathcal{Y}^{-1}(\mathcal{U}_{f*}), \mathcal{Y}(\mathcal{U}_h) = = (\mathcal{U}_{f*}, \mathcal{U}_h)$. Using this fact, the relation (3.10) and the condition (3.6), we deduce that

$$\|\mathcal{X}(\mathcal{U}_{f*+h})\|^2 - \|\mathcal{X}(\mathcal{U}_{f*})\|^2 \geqq 0 \tag{3.14}$$

QED.

For the dispersion of the o.c.f. f^* in the case of the validity of the theorem 3.1 we obtain

$$\begin{aligned} \|\mathcal{X}(\mathcal{U}_{f*})\|^2 &= \|\mathcal{Y}(\mathcal{U}_{f*}) - S\|^2 = \\ &= \|\mathcal{Y}(\mathcal{U}_{f*})\|^2 + \|S\|^2 - 2(\mathcal{Y}(\mathcal{U}_{f*}), S) \\ &= \|\mathcal{V} - \lambda\mathcal{Y}^{-1}(\mathcal{U}_{f*})\|^2 + \|S\|^2 - 2(\mathcal{Y}(\mathcal{U}_{f*}), S)\,. \end{aligned} \tag{3.15}$$

According to the definition of $\mathcal{V}$, we have

$$(\mathcal{Y}(\mathcal{U}_f), \mathcal{V}) = (\mathcal{Y}(\mathcal{U}_f), S) \tag{3.16}$$

and

$$(\mathcal{V}, S) = \|\mathcal{V}\|^2\,.$$

Using these relations we obtain from (3.15)

$$\begin{aligned} \|\mathcal{X}(\mathcal{U}_{f*})\|^2 &= \|\mathcal{V}^2\| + \lambda^2\|\mathcal{Y}^{-1}(\mathcal{U}_{f*})\|^2 - 2\lambda(\mathcal{V}, \mathcal{Y}^{-1}(\mathcal{U}_{f*})) \\ &\quad + \|S\|^2 - 2(\mathcal{Y}(\mathcal{U}_{f*}), S) = \\ &= \|\mathcal{V}^2\| + \lambda^2\|\mathcal{Y}^{-1}(\mathcal{U}_{f*})\|^2 - 2(\lambda\mathcal{Y}^{-1}(\mathcal{U}_{f*}) + \mathcal{Y}(\mathcal{U}_{f*}), \mathcal{V}) + \\ &\quad + \|S\|^2 = \\ &= \|\mathcal{V}^2\| + \lambda^2\|\mathcal{Y}^{-1}(\mathcal{U}_{f*})\|^2 - 2\|\mathcal{V}\|^2 + \|S\|^2 = \\ &= \|S\|^2 + \lambda^2\|\mathcal{Y}^{-1}(\mathcal{U}_{f*})\|^2 - \|\mathcal{V}\|^2\,. \end{aligned} \tag{3.17}$$

4.

In what follows we shall give certain conditions, which assure that the conditions of the Theorem 3.1 are satisfied.

We introduce a special class of admissible control-functions. To this purpose let us denote the denominator in the (2.7) by $B(e^{i\nu})$ and the numerator by $A(e^{i\nu})$ and define a control-function $\mathcal{U}^\lambda = \{U_n^\nu\}$, where

$$U_n^\lambda = \int_{-\pi}^{\pi} \frac{\overline{A(e^{i\nu})}\,B(e^{i\nu})}{|A(e^{i\nu})|^2 + \lambda|B(e^{i\nu})|^2}\,d\hat{V}_n(\nu)\,. \tag{4.1}$$

 From the assumption about $B(e^{iv})$ it follows that for every $\lambda > 0$

$$(4.2)\qquad \|\mathcal{U}^\lambda\|^2 = \int_{-\pi}^{\pi} \left| \frac{A(e^{iv})\, B(e^{iv})}{|A(e^{iv})|^2 + \lambda |B(e^{iv})|^2} \right|^2 \mathrm{d}\, \gamma_{\mathscr{V}}(v) < \infty$$

and then the expression (4.8) is senseful.

In the two following theorems we shall differ two possible cases.

1. The polynom $|A(z)|^2 + \lambda |B(z)|^2$ has a zero point on the unit circle for $\lambda = 0$ and for $\lambda > 0$ all zeros are within the unit circle.

2. $|A(z)|^2 + \lambda |B(z)|^2 > 0$ on and withthin the unit circle for all $\lambda \geqq 0$.

Theorem 4.1. *Let the polynom* $|A(z)|^2 + \lambda |B(z)|^2$ *fulfil the condition* 1, *then for every* $c > 0$ *there exists* $\lambda > 0$ *such that*

$$(4.3)\qquad \|\mathcal{U}^\lambda\|^2 = c$$

and

$$(4.4)\qquad \|\mathscr{X}(\mathcal{U}^\lambda)\|^2 \leqq \mathscr{X}(\mathcal{U}_f)\|^2$$

for all admissible $\mathcal{U}_f$ *for which* $\|\mathcal{U}_f\|^2 \leqq c$.

Proof. According to the assumption 1 it is

$$(4.5)\qquad \|\mathcal{U}^0\|^2 = \infty \quad \text{and} \quad \lim_{\lambda \to \infty} \|\mathcal{U}^\lambda\|^2 \to 0\,.$$

From these relations and from the continuity of $\|\mathcal{U}^\lambda\|^2$ as a function of λ it follows (4.3). To prove (4.4) it suffices to note that $\mathcal{U}^\lambda$ fulfils the conditions (3.3) and (3.4) of the theorem 3.1. Using (4.1) for (3.3) we obtain

$$(4.6)\qquad \int_{\pi-}^{\pi} \left| \frac{B(e^{iv})}{A(e^{iv})} \right|^2 \left| \frac{A(e^{iv})\, B(e^{iv})}{|A(e^{iv})|^2 + \lambda |B(e^{iv})|^2} \right|^2 \mathrm{d}\, \gamma_{\mathscr{V}}(v)\,.$$

If $|A(e^{iv})|^2 > 0$ for $v \in [-\pi, \pi]$ then the expression (4.6) is finite. $|A(e^{iv})|^2 + \lambda |B(e^{iv})|^2 > 0$ for $\lambda > 0$ according the assumption.) If $A(e^{iv}) = 0$ for some v then in such a point we can define $A(e^{iv})/A(e^{iv}) = 1$ and we obtain the same result.

In the same way it is possible to prove that the condition (3.4) holds. Then $\mathcal{U}^\lambda$ fulfills the theorem 3.1, QED.

By the same way it is possible to obtain the following result.

Theorem 4.2. *Let the polynom* $|A(z)|^2 + \lambda |B(z)|^2$ *fulfil the condition* 2, *then for every* c *for which*

$$(4.7)\qquad 0 < c < \int_{-\pi}^{\pi} \left| \frac{A(e^{iv})}{B(e^{iv})} \right|^2 \mathrm{d}\, \gamma_{\mathscr{V}}(v)$$

there exists $\lambda > 0$ *such that*

$$\|\mathscr{U}^{\lambda}\|^2 = c \tag{4.8}$$

and

$$\|\mathscr{X}(\mathscr{U}^{\lambda})\|^2 \leqq \|\mathscr{X}(\mathscr{U}_f)\|^2 \tag{4.9}$$

for all admissible $\mathscr{U}_f$ *for which* $\|\mathscr{U}_f\|^2 \leqq c$.

Before formulating the further results let us remark the following. According to our definition the random variables Y_n depend on $W_n, W_{n-1}, \ldots$ for every n. Let us not consider the limitation which follows from the fact that the process $\mathscr{Y} = \{Y\}$ is a transformation of a process $\mathscr{U}_f = \{U_n\}$ and let us admit for Y_n any function of $W_n, W_{n-1}, \ldots$, for which $E\{Y_n^2\} < \infty$. It is well known that the expression

$$\|\mathscr{X}^2\| = \|\mathscr{Y} - S\|^2 \tag{4.10}$$

has its minimal value in the class of these processes if $\mathscr{Y} = \mathscr{V}$, i.e.

$$Y_n = E\{S \mid W_n, W_{n-1}, \ldots\}\,. \tag{4.11}$$

Because the class of all processes of the considered type includes the class of the previous considered processes $\mathscr{Y}(\mathscr{U}_f)$ it is obvious that

$$\|\mathscr{V} - S\|^2 < \|\mathscr{X}(\mathscr{U}_f)\|^2 \tag{4.12}$$

and $\|\mathscr{V} - S\|^2$ is then the lower bound of the dispersion of the output.

In the case 2 we have the following theorem.

Theorem 4.3. *Let the polynom* $|A(z)|^2 + \lambda|B(z)|^2$ *fulfil the condition* 2, *then for every* $\varepsilon > 0$ *there exists* d *and* λ *such that*

$$\|\mathscr{U}^{\lambda}\|^2 = d \tag{4.13}$$

and

$$\big|\|\mathscr{V} - S\|^2 - \|\mathscr{U}^{\lambda}\|^2\big| < \varepsilon \tag{4.14}$$

where $\mathscr{U}^{\lambda}$ *is a random process given by* (4.1).

Proof. The proof of this statement is very simple. The existence of such a λ which fulfils (4.16) follows from the continuity of $\|\mathscr{U}^{\lambda}\|$ as a function of λ. This λ specifies then the corresponding d in (4.15).

Note. As $|A(z)| > 0$ for $|z| = 1$ we could put $\lambda = 0$ in (4.1), but this control-function is not admissible according to our definition. Then from the assumption that the polynom $B(z)$ is of a higher degree than $A(z)$ it follows that the control-function should depends also on the values W_m, $m > n$ for any given n. This contradicts the assumptions concerning $\mathscr{U}_f$.

(Received December 6th, 1966.)

J. L. Doob: Stochastic Processes. John Wiley & Sons, New York — Chapman & Hall, London 1953.

VÝTAH

Optimální řízení lineárního systému s náhodnými stacionárními poruchami

Otakar Šefl

Mějme systém popsaný rovnicí $x_n = y(u_n, u_{n-1}, \ldots) + s_n$, kde $\{x_n\}$ je výstup systému, $\{u_n\}$ je řídicí signál a $y(u_n, u_{n-1}, \ldots)$ je daná lineární transformace řídicího signálu. $\{s_n\}$ je stacionární náhodná porucha. Předpokládá se, že k stanovení řídicího signálu $\{u_n\}$ je možno použít jen hodnot pomocného signálu $\{w_n\}$, který je statisticky závislý na $\{u_n\}$. V článku se řeší úloha nalezení takového řízení $\{u_n\}$, které při daném omezení vlastního rozptylu minimalizuje rozptyl výstupu $\{x_n\}$.

Dr. Otakar Šefl, CSc., Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.