



Soviet-era science, translated into English

OPTIMAL LEARNING SYSTEMS

CYBERNETICS

1967

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196701.55551>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 621.398+621.52[016.3]

*CYBERNETICS
AND CONTROL THEORY*

Corresponding Member of the USSR Academy of Sciences V. S. PUGACHEV

OPTIMAL LEARNING SYSTEMS

In (1-3) a statistical theory of optimal learning algorithms for automatic systems was developed under the assumption that the probability distribution of the input signal Z and of the required output signal W is specified to within a finite number of parameters, which remain constant during learning and the subsequent operation of the system. Here a generalization of the theory of optimal learning systems is given for the case when the unknown parameters of the distribution of the input and required output signals vary randomly during learning and the subsequent operation of the system.

Suppose that all possible realizations z of the input signal Z are elements of some set A , and all possible realizations w of the required output signal W are elements of some set B . Let $\eta(\Delta z, \Delta w; \theta | \lambda)$, $\Delta z \subset A$, $\Delta w \subset B$, be a family of probability measures defined on the Cartesian product $A \times B$, depending on a parameter λ , which takes values in the set L , and on a numerical parameter θ^* . Let $\Lambda(\theta)$ be a random function of the parameter θ , taking values in the set L . We shall assume that the input and required output signals Z, W in each cycle of operation of the system are random variables distributed in $A \times B$ in accordance with the probability measure $\eta(\Delta z, \Delta w; \theta | \Lambda)$ for some value θ and for a value Λ of the parameter λ equal to the value of the random function $\Lambda(\theta)$ at the corresponding value of θ . In different cycles of operation of the system the parameter θ has different values (in particular, θ may be the number of the cycle or the moment of its beginning). Consequently, the parameter Λ has different random values in different cycles of operation of the system.

The learning of the system consists in the fact that, during N cycles corresponding to the values $\theta_1, \dots, \theta_N$ of the parameter θ , certain signals statistically related to Λ —the *learning signals*—are introduced into it. The system is required, by statistical processing of the learning signals, to produce an optimal estimate W^* of the required output signal W in the first cycle of operation after learning, ensuring fulfillment of the condition

$$M[l(W, W^* | \Lambda)] = \min, \quad (1)$$

where $l(W, W^* | \Lambda)$ is a loss function, which may depend on the unknown param-

eter Λ , and the probabilistic averaging is performed over all possible realizations of the learning signals, of the random function $\Lambda(\theta)$, and of the input, required output, and actual output signals Z, W, W^* in the first cycle after learning.

Let us denote by \tilde{W} the actual output signal of any learning system (not necessarily optimal), and by Ξ the totality of all learning signals (regarded as random variables) and of the input signal Z in the first cycle after learning. Then we shall have

* Speaking of probability measures defined on a certain set, we mean, of course, measures defined on the corresponding sigma-algebra.

$$\begin{aligned}
 R(\delta) &= M[l(W, \hat{W} \mid \Lambda)] = \\
 &= \int_L d\lambda \int_{\bar{L}} d\bar{\lambda} A(\lambda, \bar{\lambda}) \int_X d\sigma(\xi \mid \lambda, \bar{\lambda}) \int_B d\delta(\hat{w} \mid \xi) \int_B l(w, \hat{w} \mid \lambda) d\kappa(w; \theta \mid z, \lambda),
 \end{aligned} \tag{2}$$

where $A(\Delta\lambda, \Delta\bar{\lambda})$ is the joint probability measure of the random parameter $\Lambda = \Lambda(\theta)$ in the first cycle of operation of the system after training and of the aggregate $\bar{\Lambda}$ of random parameters $\Delta_1 = \Lambda(\theta_1), \dots, \Delta_N = \Lambda(\theta_N)$, $\Delta\lambda \subset L$, $\Delta\bar{\lambda} \subset \bar{L} = L^N$; $\sigma(\Delta\xi \mid \lambda, \bar{\lambda})$ is the conditional probability measure of the random variable Ξ for given values $\lambda, \bar{\lambda}$ of the random variables $\Lambda, \bar{\Lambda}$, $\Delta\xi \subset X$; $\delta(\Delta\hat{w} \mid \xi)$ is the decision function of the system in the first cycle after training, representing the conditional probability measure of its output signal \hat{W} for a given value ξ of the variable Ξ , $\Delta w \subset B$, $\kappa(\Delta w; \theta \mid z, \lambda)$ is the conditional probability measure of the required output signal W in the first cycle after training for given values z, λ of the input signal Z and the parameter Λ :

$$\kappa(\Delta w; \theta \mid z, \lambda) = \frac{d\eta(z, \Delta w; \theta \mid \lambda)}{d\gamma(z; \theta \mid \lambda)} = \frac{d\eta(z, \Delta w; \theta \mid \lambda)}{d\eta(z, B; \theta \mid \lambda)}. \tag{3}$$

In the case of self-learning, Ξ is the aggregate of input signals Z_1, \dots, Z_N, Z obtained by the system during the training period and in the first cycle after training, and $X = A^{N+1}$. In this case, if Z_1, \dots, Z_N, Z are conditionally independent for any values $\lambda, \bar{\lambda}$ of the variables $\Lambda, \bar{\Lambda}$, then the conditional probability measure of the variable Ξ is determined by the formula

$$\sigma(\Delta\xi \mid \lambda, \bar{\lambda}) = \gamma(\Delta z; \theta \mid \lambda) \prod_{i=1}^N \gamma(\Delta z_i; \theta_i \mid \lambda_i), \tag{4}$$

where $\gamma(\Delta z; \theta \mid \lambda) = \eta(\Delta z, B; \theta \mid \lambda)$ is the probability measure of the input signal Z .

In the case of training the system by a teacher by means of demonstration, (1) Ξ is the aggregate of input signals Z_1, \dots, Z_N, Z and the teacher's output

signals $\widetilde{W}_1, \dots, \widetilde{W}_N$, introduced into the system in the course of training, and $X = A^{N+1} \times \widetilde{B}^N$, where \widetilde{B} is the space of all possible realizations of the teacher's output signal \widetilde{W} (it may, in particular, coincide with B). In this case, if the triples of signals $(Z_1, W_1, \widetilde{W}_1), \dots, (Z_N, W_N, \widetilde{W}_N)$ and Z are conditionally independent for any values $\lambda, \bar{\lambda}$ of the variables $\Lambda, \bar{\Lambda}$ (W_1, \dots, W_N are the required output signals corresponding to the training input signals Z_1, \dots, Z_N), then the conditional probability measure of the random variable Ξ is determined by the formulas

$$\sigma(\Delta\xi \mid \lambda, \bar{\lambda}) = \gamma(\Delta z; \theta \mid \lambda) \prod_{i=1}^N \int_{\Delta z_i} \pi(\Delta\tilde{w}_i; \theta_i \mid z_i, \lambda_i) d\gamma(z_i; \theta_i \mid \lambda_i), \quad (5)$$

$$\pi(\Delta\tilde{w}; \theta \mid z, \lambda) = \int_B \delta_y(\Delta\tilde{w}; \theta \mid z, w, \lambda) d\mu(w; \theta \mid z, \lambda), \quad (6)$$

where $\delta_y(\Delta\tilde{w}; \theta \mid z, w, \lambda)$ is the teacher's decision function, i.e., the conditional probability measure of his output signal \widetilde{W} for given values z, w, λ of the variables $Z, W, \Lambda = \Lambda(\theta)$, $\Delta\tilde{w} \subset \widetilde{B}$.

In the case of training the system by a teacher by means of evaluating its actions, Ξ is the aggregate of training input signals Z_1, \dots, Z_N , the corresponding output signals of the system $\widehat{W}_1, \dots, \widehat{W}_N$, the teacher's evaluations (his output signals) $\widetilde{W}_1, \dots, \widetilde{W}_N$, and the input signal Z in the first cycle after training, and $X = A^{N+1} \times B^N \times \widetilde{B}^N$. In this case, if the quadruples of signals $(Z_1, W_1, \widehat{W}_1, \widetilde{W}_1), \dots, (Z_N, W_N, \widehat{W}_N, \widetilde{W}_N)$ and Z are conditionally independent for any values $\lambda, \bar{\lambda}$ of the variables $\Lambda, \bar{\Lambda}$, then the conditional probability measure of the variable Ξ is determined by the formulas

$$\begin{aligned} \sigma(\Delta\xi \mid \lambda, \bar{\lambda}) &= \gamma(\Delta z; \theta \mid \lambda) \prod_{i=1}^N \int_{\Delta z_i} \delta^i(\Delta\hat{w}_i \mid z_i) \times \\ &\times \pi(\Delta\tilde{w}_i; \theta_i \mid z_i, \hat{w}_i, \lambda_i) d\gamma(z_i; \theta_i \mid \lambda_i), \end{aligned} \quad (7)$$

$$\pi(\Delta\tilde{w}; \theta \mid z, w, \hat{w}, \lambda) = \int_B \delta_y(\Delta\tilde{w}; \theta \mid z, w, \hat{w}, \lambda) d\chi(w; \theta \mid z, \lambda), \quad (8)$$

where $\delta^1(\Delta\hat{w} \mid z), \dots, \delta^N(\Delta\hat{w} \mid z)$ are the decision functions of the system in the learning cycles, independent of the values of the parameters $\lambda_1, \dots, \lambda_N, \lambda, \Lambda(\theta_1), \dots, \Lambda(\theta_N), \Lambda(\theta)$.

Representing (2) in the form

$$R(\delta) = \int_X d\beta(\xi) \int_B \rho(\xi, \hat{w}) d\delta(\hat{w} \mid \xi), \quad (9)$$

$$\rho(\xi, \hat{w}) = \int_L d\Omega(\lambda | \xi) \int_B l(w, \hat{w} | \lambda) d\chi(w; \theta | z, \lambda), \quad (10)$$

where

$$\beta(\Delta\xi) = \int_L d\lambda \int_{\bar{L}} \sigma(\Delta\xi | \lambda, \bar{\lambda}) d_{\bar{\lambda}} A(\lambda, \bar{\lambda}) \quad (11)$$

is the unconditional probability measure of the random variable Ξ , and

$$\Omega(\Delta\lambda | \xi) = \int_{\Delta\lambda} d\lambda \int_{\bar{L}} \frac{d\sigma(\xi | \lambda, \bar{\lambda})}{d\beta(\xi)} d_{\bar{\lambda}} A(\lambda, \bar{\lambda}) \quad (12)$$

is the conditional (posterior) probability measure of the parameter $\Lambda = \Lambda(\theta)$ for the given realization ξ of the random variable Ξ , and assuming that for every realization ξ of the random variable Ξ there exists a set $C_\xi \subset B$ of values w^* of the random variable \hat{w} satisfying the condition

$$\rho(\xi, w^*) \leq \rho(\xi, \hat{w}) \quad \text{for all } \hat{w} \in B, \quad (13)$$

we arrive at the conclusion that the optimal decision function of the learning system $\delta_{\text{opt}}(\Delta\hat{w} | \xi)$ will be an arbitrary conditional probability measure completely concentrated on the set C_ξ . In the particular case when, for every ξ , there exists a unique value $w^* = B\xi$ (B is a deterministic operator) satisfying (13), the algorithm of the optimal learning system proves to be deterministic.

A learning system that produces its output signal W^* (the optimal estimate of the required output signal W) in the first cycle after learning by minimizing expression (10) is a **Bayes optimal learning system**. No learning system that has received the same collection of realizations of the learning signals and the same realization of the input signal in the first cycle after learning can be better than the Bayes optimal learning system, i.e., can give a better estimate of the required output signal in the first cycle after learning. Consequently, the quality of the Bayes optimal learning system, determined by the value $R(\delta_{\text{opt}})$, can serve as a characteristic of the limiting possibilities of learning—the **potential learnability** of automatic systems.

The basic property of the Bayes optimal learning system, as is clear from (4), (5), (7), (11), and (12), is that, in order to determine the posterior distribution of the parameter Λ , it uses not only the realizations of the learning signals but also the realization of the input signal obtained after learning. In other words, the Bayes optimal learning system continues to improve after learning by means of self-

self-training in the course of operation. In the particular case when there is no training period, such a system continuously computes, during the operating

cycle, the a posteriori distribution of the parameter Λ from the segment of the realization of the input signal obtained up to the given moment t , and uses this distribution in producing the value of its output signal at the moment t , i.e., it carries out self-training. In this particular case the Bayesian optimal learning system is a **Bayesian optimal adaptive (self-adjusting) system**.

To evaluate the effectiveness of the optimal training process it is necessary to compare the quantity $R(\delta_{\text{opt}})$, corresponding to the Bayesian optimal learning system, with the quantity $R(\delta_\lambda)$, attainable with the aid of a **Bayesian optimal system with complete information about the parameter** Λ , producing its output signal W_λ^* by minimizing the expression

$$\rho_\lambda(\xi, \hat{w}) = \int_B l(w, \hat{w} | \lambda) d\mathcal{N}(w; \theta | z, \lambda) \quad (14)$$

for the exactly known value λ of the parameter $\Lambda = \Lambda(\theta)$. Since always $R(\delta_{\text{opt}}) \geq R(\delta_\lambda)$, when $R(\delta_\lambda) > 0$ the quantity

$$\varepsilon(\delta_{\text{opt}}) = \frac{R(\delta_{\text{opt}}) - R(\delta_\lambda)}{R(\delta_\lambda)} \quad (15)$$

may serve as a measure of the limiting attainable (potential) closeness of the learning system to the optimal system with complete information about the distribution of the input and required output signals Z, W (3).

From the formulas obtained, which determine the a posteriori measure $\Omega(\Delta_\lambda | \xi)$, it is seen that in the considerably more general learning model studied here than in (1-3), all the basic properties of optimal training processes derived in (1-3) are preserved, except for the recurrence of the process of computing $\Omega(\Delta_\lambda | \xi)$. The latter property is preserved only if the random function $\Lambda(\theta)$ is a Markov random process, i.e., if the values $\Lambda(\theta_1), \dots, \Lambda(\theta_N), \Lambda(\theta)$ of the parameter Λ in the various cycles form a simple Markov chain.

The theory of optimal training processes for automatic systems, developed in (1-3), is obtained as a particular case of the general theory set forth here, when L is a finite-dimensional Euclidean space, the signals Z, W, \hat{W} , and \bar{W} are finite-dimensional random vectors or vector random functions, and the random parameter Λ retains a constant value during the training process and after training,

$$\Lambda = \Lambda(\theta_1) = \dots = \Lambda(\theta_N) = \Lambda(\theta).$$

Received
17 V 1966

CITED LITERATURE

1. V. S. Pugachev, DAN, **172**, No. 5 (1967).
2. V. S. Pugachev, Report at the Third All-Union Conference on Automatic Control (Technical Cybernetics), Odessa, 20-25 IX 1965.
3. V. S. Pugachev, Report at the Third IFAC International Congress on Automatic Control, London, 20-25 VII 1966.

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

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.