DATA DESCRIPTION OF A SYSTEM

Pavel NEVŘIVA
Technical University of Ostrava
17. listopadu
708 00 Ostrava
Czech Republic

Keywords:
Adaptive Algorithms, Accumulation of Data, Statistical Prediction, Information

Abstract

In this paper, a brief discussion on description of process by memorized data is given. The insight into the problem can offer modified views on optimal control, on data compression at communication systems with respect to information content of message, etc.

The idea of process description by memorized data with different information content will be presented here on the classical case study of optimal control: the data based control algorithm (data algorithm, DA) gathers data from the controlled process and derives control signal (control) from data accumulated in the data base. The implementation of the DA on the ideal computer which is not limited by its speed or capacity of memory is expected for simplicity. Accuracy of the data algorithm is then given by a-priori knowledge of the task and by information exchange between the controlled process and the computer.

List of Symbols

k (l)	time index, k = 1, 2, ...
t	time
$\#_k$	$\#(t_k)$
i (j,m,n)	space index, i = 1, 2, ...
s	state of the controlled system
${}^{i}s$	i-state of the controlled system
$\mathbf{s}$	{ s }
S	state space of the controlled system
v	state of the external area
${}^{i}v$	i-state of the external area
$\mathbf{v}$	{ v }
V	state space of the external area
x	state of the controlled process
${}^{i}x$	i-state of the controlled process
$\mathbf{x}$	{ x }
X	state space of the controlled process, X = S + V
w	control
${}^{i}w$	optimal control from the state ${}^{i}x$
${}^{ij}w$	optimal control transferring the system from the state ${}^{i}x$ = { ${}^{i}s$, ${}^{i}v$ } to { ${}^{j}s$, v }
W	state space of control
q	cost index
${}^{ij}q$	cost index increment corresponding to the transfer of the process from the state ${}^{i}x$ = { ${}^{i}s$, ${}^{i}v$ } to the state { ${}^{j}s$, v }
r	values of the additional collected parameters
y	y = (w , q , r , t)
$\mathbf{y}$	{ y }
Y	space of the DA
z	z = (x, y)
$\mathbf{z}$	{ z }
Z	Z = X + Y

1. Introduction

The call for the physical systems prediction/control has led to the models describing the particular system as the unity, using the standard mathematical functions. Such models will be mentioned here as the global models.

The global models have been developed from observations which had been, the exception proves the rule, of local character. The sets of data describing systems in discrete points of its state space will be introduced here as the local models and their implementation on the computer as the data models. The basic tasks of the data model and the global model are identical: to reflect the causal relations between state variables changes. Discussing the control system the model has to procure the prediction of the optimal control. In the classic procedure the a-priori hypotheses dealing with

- physical interpretation of approximating function
- physical interpretation of quantification of measurement errors
- physical interpretation of measurement errors distribution
- location of nonregular measured data

are used.

Being the above hypotheses in accordance with reality, they carry information usable for interpretation of the measurement, in the opposite case they represent an interference. In addition to these physical conditions a practical condition is present: - the reason for the global model exists.

The need of the explicit function $y = y(x)$ is less frequent then the claim for y_i corresponding the given x_i in applications.

Local/data models will not be suitable for situations where the functions, describing the system have the physical basis. In the situations where the global models are supported mostly by historical arguments local models offer alternative procedures. Such cases are frequent in the application area and it may be useful to go into the problem.

2. Data Models

State space of the controlled process

$$\mathbf{X} = \mathbf{S} + \mathbf{V}$$

where **S** is the state space of the controlled system
V is the state space of the external area

It is given by the set { x } = { s, v } of discrete states. We will not index particular states x in the text, the discussion will be led for any $x \in \mathbf{X}$.

To control the system the x_k is measured in t_k and w_k is predicted by the DA from memorized z_k. Subsequently the values of w_k, q_k, r_k are measured and the data $y_k = (w_k, q_k, r_k, t_k)$ are stored.

The interval of the state space sampling has to be sufficiently small and frequency of measurement has to be sufficient; as well as the stock of collected data. The set of retained data is z, it presents the data description of the process. The set z changes within the time. At the time t_k the DA can predict w_k from the set z_k (which does not contain y_k). Evidently, in practical application of the DA the algorithm has to incorporate procedures of initial assessment of the control w to the measured state x. These procedures can be derived from a-priori familiarity with the system, such procedures are always used for checking the boundaries of control, for nonstandard control actions, etc.

Let us assume an initial connection w to x. Using a method of dynamic programming, algorithm predicts the optimal trajectory from a measured state ${}^i x$ to the desired state and searches the time series of control, which transfer the system to this final state. At t_k it looks for ${}^{ij}w_k$ which transfers the system from ${}^i x_k$ to the neighbouring state ${}^j x$ on the optimal trajectory. Let us start the discussion for the algorithm which computes the accurate correction Dw_k to the optimal w_k after the local control action has been done by predicted w_k, $Dw_k \in r_k$. This demand is too strict for practice, where only the knowledge of the sign of Dw_k is often wanted; this may be possibly also the easiest requirement of the correlation between the theory of automatic control and the a-priori knowledge of the physical essence of the controlled system.

The DA imitates the behaviour of a person, who for his decisions consults the memorized experience.

An advantage of the DA is that the local model of system can be often assumed to be in its local states x linear, stationary and ergodic also at system whose global model is nonlinear, nonstationary and nonergodic. Disadvantage of the DA concept lies in large requirements on technology, computer, for implementation of a DA. DA can be regarded as an extreme type of classical control algorithm, fuzzy algorithm or neuron net control algorithm. Time series $\{{}^{ij}w\}$ and corresponding time series $\{{}^{ij}q\}$ correspond to different types of local models:

Stationary deterministic model. The ${}^{ij}q$, ${}^{ij}w$ are constant. It is possible to map the state space by standard calculation, the map is time invariable. For the fixed setpoint it is possible to assign the optimal control w to any x, the optimal trajectories are constant. The feedback serves for discretisation and disturbance compensation.

Nonstationary deterministic model. The ${}^{ij}q$, ${}^{ij}w$ are deterministic. It is possible to map the state space by standard calculation, the map is time variable. It is necessary to set the optimal trajectory at given state of the system. The feedback serves for discretisation and disturbance compensation. The optimal control of a system with moving obstacle in the state space was published in [4].

Stationary stochastic model will be briefly discussed further on. It enables profitable use of accumulated data and suits for the DA. At stationary stochastic process it is not possible to reduce data without loss of information [1], [2]. In case of the lack of a-priori knowledge of the controlled system, the discussion of the control process can be led only on the basis of measured data.

Nonstationary stochastic model requires both the continuous data logging and the continuous computation of control vector. Considering that the data model operates with all procurable data it is theoretically possible to design the DA which will work as well as any other one.

3. Validity of Accumulated Data

A control process described by a stationary stochastic model is studied. A discussion will be restricted on the prediction w_l of the optimal control w_l from the given w_k; $w_k = w_k + Dw_k$. The probability density of w_l at $t_l > t_k$ determines the chance of predicting an useful w_l from w_k. For the continuous time and the Gaussian random variable w(t) the best estimate w_l of w_l is

$$w_l = w_k \cdot \frac{f(Dt)}{f(0)} \qquad (1)$$

with the best estimate σ_l of standard deviation σ_l

$$\sigma_1 = \sqrt{f(0) \cdot \left[1 - \left(\frac{f(Dt)}{f(0)} \right)^2 \right]} \qquad (2)$$

where

$$f(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} w(t-\tau) \cdot w(t)\, dt \qquad (3)$$

is the autocorrelation function of w(t), $Dt = t_l - t_k$.
For $Dt \to 0$ the best estimate of control is $w_l = w_k$, with the best estimate of standard deviation $\sigma = 0$.
For $Dt \to \infty$ the best estimate of control is $w_l = w_{med}$, with the best estimate of standard deviation $\sigma = \sigma$.

The prediction of ${}^{i,j}w_l$ from ${}^{m,n}w_k$ using the correlation function

$$g(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T m,n} {}^{i,j}w(t).w(t-\tau)\, dt \qquad (4)$$

where $({}^{i}s, {}^{m}s)$ and $({}^{j}s, {}^{n}s)$ are couples of close states is meaningful. Estimated values of w can be weighted by its accuracy defined by standard deviation. The estimate of optimal w from the low frequented state can be found out in the area surrounding this state in the state space.

Continue the discussion for the algorithm that can compute only the medium value $\mathbf{Dw}_k$ of the correction Dw_k after the local control action has been done by w_k. Dw_k is given by $\mathbf{Dw}_k$ and by the standard deviation σ_k. The best estimate of control w_l is then given by (1) where $(w_k + \mathbf{Dw}_k)$ is substituted for w_k. The best standard deviation estimate σ_l is given by

$$\sigma_1 = \sqrt{f(0) \cdot \left[1 - \left(\frac{f(Dt)}{f(0)} \right)^2 + \sigma_k^2 \cdot \left(\frac{f(Dt)}{f(0)} \right)^2 \right]} \qquad (5)$$

Statistical prediction described has its limits in the description of stochastic functions by second order moments. For different premises the different attempts would be effective. The analysis of these problems turns de-facto into analysis of the a-priori knowledge of the control process.

Equations (1) to (5) point to a limited use of both the infinite speed and infinite memory capacity of an ideal computer.

4. Information Provided by Measured Data

The interpretation of Shannon's idea on a measure of the information provided by an experiment has been given by Lindley in the very early paper [3]. The amount of information I provided by measured value $\alpha \in \mathbf{A}$ for parameter $\beta \in \underline{\beta}$ estimation has been defined by Lindley as $I = I_0 - I_1$ (6)

where

$$I_0 = -\int_{\underline{\beta}} p(\beta) \cdot \log_2 p(\beta)\, d\beta \qquad (7)$$

$$I_1 = -\int_{\underline{\beta}} p(\beta|\alpha) \cdot \log_2 p(\beta|\alpha)\, d\beta \qquad (8)$$

where $p(\beta)$ is the a-priori probability density function of β (before α is measured)
$p(\alpha|\beta)$ is the (a-posteriori) conditional probability function of β (after α has been measured)

In comparison with the problem solved in [3] the control system is a dynamic one under the influence of control and disturbances. At classical control systems where w(t) is derived from the equations describing adaptive controller the discussion of the amount of information provided by feedback leads to discussion of a-priori and a-posteriori probability density function of their parameters.

The DA optimizes w_k locally at single points of x, integrands in (7) and (8) contain probability density function and conditional probability density function of control w_k. These in contradistinction to [3] are not given and have to be approximated by a-priori and a-posteriori statistical estimate. Then

$$DI_l = I_l - I_{l-1} \qquad (9)$$

where
$$I_{l-1} = -\int_W p(w_l|z_{l-1}) \cdot \log p(w_l|z_{l-1})\, dw \qquad (10)$$
$$I_l = -\int_W p(w_l|z_l) \cdot \log p(w_l|z_l)\, dw \qquad (11)$$

where $p(w_l | z_{l-1})$ is the a-priori probability density function of w_l estimate (before z_l is measured at t_l)
$p(w_l | z_l)$ is the a-posteriori probability density function of w_l estimate (after z_l has been measured at t_l)

The DA derives w_l from the set z_l :

- process without control: $\mathbf{Z} = \mathbf{0}$ (12)
- adaptive control without feedback: $\mathbf{Z} = \mathbf{V} + \mathbf{Y}$ (13)
- adaptive control: $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ (14)
- etc.

The computer capacity can be apportioned among different tasks with respect to the information flow amount. Algorithm simulates the behaviour of a living creature being able to analyse unpredicted local changes of a very large scene observed. There is neither a procedure to set $p(w | z)$, nor the procedure to set w from $p(w | z)$ in (9) to (11). More, there is not inherent relation between information amount and the quality of control. It streams from the fact that Shannon's information is not connected with verity.

5. Conclusion

The review of some problems associated with the idea of algorithm operating with the ideal computer and the real system has been given in this paper. At case study, the close connection of the quality of control and the a-priori knowledge was presented, ideal computer is not sufficient condition for the optimal control.

At data systems, the data exchange between both computer and system as well as computer and network can be extremely reduced with respect to information associated with the data. Ideal computer is an idea for the future, there are not systems running on such a computer. On the other hand, some slow systems of today are a very good approximation of these ones.

6. Rererences:

[1] GRAY R.M.: Entropy and Information Theory. Springer-Verlag, Berlin 1992.

[2] KULLBACK, S. AND LEIBLER, R.A. : Information and sufficiency. Ann. Math. Stat., Vol. 22 (1951), pp.79-86.

[3] LINDLEY, D.V. : On a Measure of the Information Provided by an Experiment. Ann. Math. Stat., Vol. 27 (1956), pp. 986-1005.

[4] NEVŘIVA P.: Regulátor, využívající paměťovou kapacitu počítače. Automatizace 27, 1984, c.11. FMHTS Praha, Praha 1984.

[5] NEVŘIVA P.: Data-based Optimal Control Algorithm. 2nd International Symposium on System Analysis and Simulation. AV NDR. Berlin 1985.

[6] PREUSS H.P., Tresp V. : Neuro-Fuzzy. Automatisierungstechnische Praxis 36 (1994) No.5. R. Oldenbourg Verlag, Berlin 1994.

About author ...

Pavel NEVŘIVA received the MSEE degree (1963) and the Ph.D. degree (1968) in Cybernetics from Technical University of Brno and the DrSc degree (1986) in Technical Cybernetics from Czech Technical University of Prague. His research interests include information processing and automatic control of large processes.

RADIOENGINEERING REVIEWERS

April 1996, Volume 5, Number 1