ON OPTIMALITY CRITERIA IN IDENTIFICATION PROBLEMS

I.S. Durgarian N.S. Rajbman
Institute of Control Sciences
Moscow, USSR

The identification problem is largely formulated as a problem of finding a plant operator estimate optimal in terms of a specified criterion. The resultant solution permits estimation of the model adequacy. The criterion and the model structure are selected individually for each model. For complex systems the large number of plant parameter interrelations and of external disturbances make a completely adequate model non-feasible or undesirable. True, a more accurate model and a better system description is supposed to give a better forecast and control of the system but studying each input and its responses takes more time, funds and material resources. Hence it is clear that identification of large systems requires quantitative assessment of the effect of each system input on its output variable and of the accuracy with which a model where these inputs are integrated simulates the actual processes; a decision should be made, which system variables should be represented in the model. Identification is also impossible without systems approach, studying the hierarchical structure and knowledge of the entire system functioning.

We will describe optimal selection of plant information indices with the techno-economic criterion. The results related to the model performance act as constraints.

1. The desired product accuracy for a complex plant or especially for a set of plants can be obtained in a number of ways, for instance, by varying the input or state variables or both. Calculation of optimal characteristics requires establishment of input and state variable indices such that would characterize the internal state of plants and ensure the desired output quality in a "best" way, or by a specified criterion, or objective function. Model referenced solution generally relies on numerical characteristics closely related to specified requirements. The mean output variable characterizes the nominal value of a qualitative index (centers middle of the tolerance field, nominal size, etc.); variance, the

admissible deflection of the output variable (tolerance field); entropy, the output variable scatter. Consequently, the control should ensure the desired values of output and state numerical indices. Quite naturally, methods of ensuring the desired quality vary with optimality criteria and a control optimal in terms of one criterion can be far from the optimum in terms of the other. Let us consider certain details of techno-economic optimal control and try meeting the requirements of a certain comprehensive index integrating several techno-economic indices. Selection of optimality criteria is a major complex affair untractable by purely mathematical tools. Unlike the statistically optimal systems which, if optimal in terms of one criterion, are near-optimal in terms of others, these systems are not optimal in terms of another criterion.Therefore selection for a complex industrial process is heavily dependent on specific conditions and problems posed in design of new/or automation of existing processes. In many cases the definition of general design problems includes optimality criteria or the necessity of comparative analysis of data obtained with different criteria is indicated.

This section will describe the technique of calculating the basic indices of a technological line so as to ensure the desired production at minimal costs.

Let the technological chain be modelled as multiple regression of the output variable with respect to input variables

$$M\{y(t)/X_1(t),\ldots,X_n(t)\} = \sum_{i=1}^{n} b_i X_i(t). \qquad (2.1)$$

For dynamic plants the regression integrates the lag τ with respect to $X_i(t-\tau_i)$. In further discussion the arguments are not written so as not to encumber the equations.

The variance at the output can be given as a sum of two variables

$$\mathcal{D}\{y\} = \mathcal{D}\{M\{y/x_1,\ldots,x_n\}\} + M\{\mathcal{D}\{y/x_1,\ldots,x_n\}\}, \qquad (2.2)$$

where $\mathcal{D}\{M\{y/X_1,\ldots,X_n\}\}$ characterizes that part of the overall variance for the output variable y which is caused by the inputs and $M\{\mathcal{D}\{y/X_1,\ldots,X_n\}\}$ is that part which is caused the by other factors except $X_1,\ldots,X_n$.

The following formula for $\mathcal{D}\{M\{y/X_1,\ldots,X_n\}\}$ is convenient for practical computation

$$\mathcal{D}\{M\{y/x_1,\dots,x_n\}\} = \sum_{i=1}^{n} b_i \operatorname{cov}(x_i\, y), \tag{2.3}$$

where $\operatorname{cov}(x_i\, y) = M\{[x_i - M\{x_i\}][y - M\{y\}]\} = r_{x_i y}\sigma_i\sigma_y .$

The output variable found through (2.1) should ensure the conditions

$$\mathcal{D}\{M\{y/x_1,\dots,x_n\}\} \leq \mathcal{D}_3\{y\}, \tag{2.4}$$

where $\mathcal{D}_3\{y\}$ is the specified value of the output variable variance.

For an automatic line of n jobs the total cost of an article or part C_Λ is composed of the input quality cost and the cost of each of the operations C_i

$$C_\Lambda = \sum_{i=0}^{n} C_i \quad . \tag{2.5}$$

The costs C_i (in this case techno-economic indices are used as optimality criteria) can be represented in the form

$$C_i = A_i + f(\delta_i), \tag{2.6}$$

where A_i are constant values of the index elements independent of the accuracy δ_i for the output product and $f(\delta_i)$ are variable values of index elements dependent on δ_i .

The shape of the dependence $f(\delta_i)$ can normally be determined for a specific plant or line on the knowledge of normal operation data.

If the accuracy is characterized in terms of the r.m.s. error σ then formula (2.6) can be written in the form

$$C_i = A_i + \frac{k_i B_i}{6\sigma_i} \tag{2.7}$$

where the values A_i and B_i are determined for each job and k_i is the coefficient of relative scatter whose value depends on the job error distribution law. In more precise terms the derivation of (2.7) is given in Ref. 1. That formula represents the trade-offs of the job cost C_i and the job output accuracy characteristic σ_i .

Knowing the line characteristics and their relations to techno-economic indices one can compute the optimal line by mathematical programming, resolving multipliers or Lagrange conditional multiplier techniques.

Let us formulate the following nonlinear programming problem : Find the minimal value of the overall machining cost

$$\gamma = \sum_{i=0}^{n} \frac{k_i B_i}{6\sigma_i} \tag{2.8}$$

defined as a function of n variables $\sigma_1, \dots, \sigma_n$ which should satisfy the constraints

$$\sum_{i=1}^{n} b_i \tau_{x_i y} \sigma_i \sigma_y \leq \mathcal{D}_3 \{y\}$$
$$\sigma_{i\,max} > \sigma_i > \sigma_{i\,min} \tag{2.9}$$

If the trade-offs of the costs and the r.m.s. deflection are linear and (2.8) can be given in the form

$$k = \sum_{i=1}^{n} p_i \sigma_i , \tag{2.10}$$

where p_i are certain coefficients then we arrive at a linear programming problem with the same constraints (2.9).

In the case when the linear function (2.10) is maximal inside the definition region and the constraint (2.4) is given as the equality

$$\sum_{i=1}^{n} c_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_i c_j \tau_{x_i x_j} \sigma_i \sigma_j = \mathcal{D}_3 \tag{2.11}$$

finding the conditional extremum of the function is reduced to finding the extremum of the function

$$\Phi = K + f\beta, \tag{2.12}$$

where f is a Lagrangean multiplier and

$$\beta = \sum_{i=1}^{n} c_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_i c_j \tau_{x_i x_j} \sigma_i \sigma_j - \mathcal{D}_3 = 0$$

The vanishing of the partial derivatives leads to equations for unknown σ_i . For the case under consideration we have

$$\sigma_i^2 = \frac{p_i^2 \, \mathcal{D}\{M\{y/x_1,\dots,x_n\}\}}{4c_i^2\left[\frac{1}{4}\sum_{i=1}^{n}\frac{p_i^2}{c_i^2} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=i+1}^{n} r_{x_i x_j}\frac{p_i p_j}{c_i c_j}\right]} \tag{2.13}$$

If the value of γ in eq. (2.8) is to be minimized with

$$\varphi = \sum_{i=1}^{n} b_i r_{x_i y}\sigma_i\sigma_y - \mathcal{D}_3 = 0,$$

the Lagrange method leads to the expression

$$\sigma_i = \frac{\mathcal{D}\{M\{y/x_1,\dots,x_n\}\}\left(k_i B_i/6\sigma_y b_i r_{x_i y}\right)^{1/2}}{2{,}45\sum_{i=1}^{n}\left(k_i B_i \sigma_y b_i r_{x_i y}\right)^{1/2}} \tag{2.14}$$

Similarly the problem of finding the plant optimal entropic characteristic of is solved when the accuracy of plant functioning depends on the entropy of its output variable [2].

Note that in equality (2.2) the first sum represents the model variance, i.e. $\mathcal{D}\{y^*\}$. Similarly, represent the entropy of the plant output in the form:

$$H\{y\} = H\{y^*\} + \Delta H \tag{2.15}$$

where ΔH represents an error occurring when the plant is replaced by its model.

The linear model of the plant with n inputs and n outputs can be described in the form

$$y^* = AX$$

where $X = (X_1,\dots,X_n)$ is the vector of model inputs,
$y^* = (y_1^*,\dots,y_n^*)$ is the vector of model outputs, and

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$ is a

matrix of regression coefficients.

Then the entropy $H\{y^*\}$ can be expressed in terms of the entropy $H\{X\}$ in the following way[2]:

$$H\{y^*\} = H\{X\} + \log \frac{\ell_x}{\ell_{y^*}} |\det A| \quad , \tag{2.16}$$

where ℓ_x and ℓ_{y^*} are degrees of accuracies of measuring the values X and y^*, and $|\det A|$ is the magnitude of the determinant of the matrix A.

Formula (2.16) is correct in the case of a linear model of the plant. However in a more general case, where for vectorial X and y^* $y^* = \varphi(X)$
in this case there exists the unambiguous inverse transformation $X = \psi(y^*)$ we have

$$H\{y^*\} = H\{X\} + M\left[\log |\mathcal{J}(X)|\right] + \log \frac{\ell_x}{\ell_{y^*}} , \tag{2.17}$$

where $\mathcal{J}(X)$ is the Jacobian of the function φ, turning into its derivative in the case of the scalar X and y^*. It is obvious that in the linear case

$$\mathcal{J}(X) = \det A$$

we shall also arrive at the expression (2.16).

In the general case the entropy of the vector X can be expressed in terms of entropies of components as

$$H\{X\} = \sum_{i=1}^{n} H\{X_i\} - \sum_{i=2}^{n} I(X_1, \ldots, X_{i-1} \; ; \; X_i), \tag{2.18}$$

where $I(X_1, \ldots, X_{i-1} \; ; \; X_i)$ is the amount of data on the value X_i contained in the vector $(X_1, \ldots, X_{i-1})$.

In the simplest case where the joint distribution of the values $X_1, \ldots, X_n$ is expressed by the normal law (note that this does not testify to the normality of the joint distribution (X, Y), or linearity of the regression), (2.18) can be written in the form

$$H\{X\} = \sum_{i=1}^{n} H\{X_i\} + \sum_{i=2}^{n} \log \left(1 - R^2(X_i \; ; X_1, \ldots, X_{i-1})\right)^{1/2} \tag{2.19}$$

where $R(X_i\,;\,X_1,\dots,X_{i-1})$ is a multiple coefficient of correlation X_i with $(X_1,\dots,X_{i-1})$.

Substituting (2.19) into (2.17) we have

$$H\{y^*\} = \sum_{i=1}^{n} H\{X_i\} + \sum_{i=2}^{n} \log\left(1 - R^2(X_i\,;X_1,\dots,X_{i-1})\right)^{1/2} + + M\left[\log|J(X)|\right] + \log\frac{l_x}{l_{y^*}} \qquad (2.20)$$

Assume that the entropy $H\{y^*\}$ should be maintained at a certain given level. The necessary entropy $H\{X\}$ of the input can be obtained immediately by equations (2.17) or (2.20). In more complicated cases it is necessary either to investigate the interrelations between the entropies of input values or to introduce additional constraints. Determine what the values $H\{X_1\},\dots,H\{X_n\}$ should be to provide the necessary value $H\{y^*\}$.

With some assumptions following from the physical nature of the plants the following dependence of the C_i on the input entropy H_i can be assumed for the i-th input:

$$C_i = \frac{1}{K_i}\log\frac{H_{oi}}{H_i},$$

where H_{oi} is the initial entropy of the i-th input, K_i is a coefficient. The values H_{oi} and K_i are considered to be given. For simplicity assume $K_1 = \dots = K_n = K$.

Formulate now the following problem. It is desired to find the values $H_1,\dots,H_n$ satisfying the equation:

$$\sum_{i=1}^{n} H_i - H\{y^*\} + \sum_{i=2}^{n} \log\left(1 - R^2(X_1,\dots,X_{i-1}\,;X_i)\right)^{1/2} + + M\left[\log|J(X)|\right] + \log\frac{l_x}{l_{y^*}} = 0, \qquad (2.21)$$

so that the total costs

$$\Phi = \sum_{i=1}^{n} C_i = \frac{1}{K}\sum_{i=1}^{n}\log H_{oi} - \frac{1}{K}\sum_{i=1}^{n}\log H_i \qquad (2.22)$$

should be minimal.

The conditional extremum of the expression Φ is obtained when the values of the input entropies are equal, i.e. when

$$H\{X_1\} = \dots = H\{X_n\} = \frac{1}{n}\sum_{i=1}^{n} H\{X_i\}.$$

If we do not restrict ourselves to the case of equality of all k_i then we obtain the extremum when the values are related as follows:

$$H\{X_1\} \cdot k_1 = H\{X_2\} \cdot k_2 = \ldots = H\{X_n\} \cdot k_n .$$

3. A most complete solution of the optimization problem with a techno-economic criterion would require several stages: 1) data acquisition for control; 2) data processing and transfer; 3) decisions-making using the results of data processing and forming control in accordance with the decision made.

Usually the predicted accuracy of plant functioning can be maintained by means of various alternative sequences of actions at each stage. Completion of each stage of work involves certain expenditures. Denote rosts for data acquisition as a; decision - making and control generation costs as c ; profit (as a result of increasing the accuracy of plant functioning) as d .

It is obvious that the expenditures a, b, c and d are functions of the input X chosen in identification (in a general case of the vectorial input.). Moreover, assume a, b, c and d are functions of some parameters α, β, γ . These can be associated with different methods and techniques of data acquisition and processing and ways of their utilization etc.

In the light of the above, the maximal profit in the case of utilization of our model can be found by the formula

$$P = \max_{\alpha,\beta,\gamma} \{ d(X,\alpha,\beta,\gamma) - a(X,\alpha) - b(X,\beta) - c(X,\gamma) \} = \\ = y(X) - z(X) . \tag{3.1}$$

The physical sense of (3.1) is obvious. The profit is a function of the income $y(X)$, obtained in control, and costs $z(X)$ of this control.

It can be assumed that the values y and z are related as follows:

$$y = y_m (1 - c e^{-kz}) , \tag{3.2}$$

where y_m is income obtained in the case of "ideal" control, i.e. when the realizable value is maintained exactly at a predicted level; c and k are coefficients the values of which are determined for process (a plant). Considering (3.2), formula (3.1) can be rewritten in the form

$$p = y(z) - z = y_m(1 - ce^{-kz}) - z = p(z) . \tag{3.3}$$

The value of the expenditure z_0 at which the maximum is achieved in the expression (3.3) represents those costs at which the maximal profit will be provided equal to

$$Q = \max_X \Big\{ \max_{\alpha,\beta,\gamma} \big\{ d(X,\alpha,\beta,\gamma) - a(X,\alpha) - b(X,\beta) - c(X,\gamma) \big\} \Big\} = \max_z p(z) = p(z_0) . \tag{3.4}$$

For finding z_0 it is sufficient to differentiate (3.3) over z and solve the equation thus obtained

$$y_m kc e^{-kz} - 1 = 0 . \tag{3.5}$$

It is obvious that the value $p(z)$ can be used as a criterion of identification and control performance. In the best variant the value $p(z)$ is maximal and equal to (3.4).

<u>References</u>

I. Райбман Н.С. Корреляционные методы определения приближенных характеристик автоматических линий. Изд-во АН СССР, "Энергетика и автоматика", №I, I96I.

2. Пугачев В.С. Теория случайных функций и ее применение к задачам автоматического управления. Физматгиз, I962.