

STRUCTURE ERRORS IN SYSTEM IDENTIFICATION*

G. A. Bekey and F. Y. Hadaegh
Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, California 90089-0781

1. INTRODUCTION

System identification is concerned with the determination of a model whose behavior approximates that of a given physical system as closely as possible, under approximately restricted experimental conditions. In practice, linear system identification is often separated into two parts: (a) determination of the order of the optimum linear model, and (b) estimation of the parameter values of the resulting linear model. Clearly, in a linear system, the model structure is determined by the choice of order. However, when the system is not linear and the nonlinearities are either omitted or incorrectly represented in the model, erroneous estimates of parameters may be obtained. This problem, i.e., the effect of erroneous assumptions of model structures on parameter values has received inadequate attention in the past. A number of books (e.g., [1]) describe variety of algorithms for system identification. The effect of erroneous assumptions of model structure generally shows up in the covariance matrix of the estimated parameters. However, this is at best an indirect indication, since an increase in the variability of the estimated parameters may also be due to a neglect of time variations and other factors. Incorrect structure assumptions may also manifest themselves in the goodness of fit criteria by which the quality of the model is judged. Thus, an incorrect structural assumption may produce a worse agreement between model outputs and system outputs. It is important to note that this is not always the case since it may be possible for the identification algorithm to select incorrect parameter values in order to compensate for erroneous assumptions of structure. This paper is concerned with an approach to system identification which explicitly takes structure errors into account and hence provides a systematic way for answering questions concerning the magnitude of estimated parameter errors resulting from structural errors.

2. FORMULATION OF THE PARAMETER IDENTIFICATION PROBLEM

Assume that there exists a physical process (the system) with inputs $u(t)$ and outputs $y_p(p)$, which are measurable. We characterize the process by the assumed mathematical model

$$\dot{x}_m(\theta, t) = f(x_m(\theta, t), u(t), x_0, \theta, t) \quad (1)$$

* This research was supported by the National Science Foundation under Grant No. NSF 8200882.

and

$$y_m(\theta, t) = g(x_m(\theta, t), u(t), x_0, \theta, t) \quad (2)$$

where the dimensions of the vectors x , y , u and θ are n , m , r and q respectively. x_m represents the state of the model while y_m are the model outputs. The function $f(\cdot)$ represents our assumption about the structure of the process, while the function $g(\cdot)$ represents the measurement operations. Equations (1) and (2) are in fact a class of models which is parameterized by the vector θ . Hence, the process of parameter identification leads to a selection of a member of this set of models, on the basis of observations of inputs and outputs of both process and model.

In the real world our assumption of model structure is never in complete agreement with that of the process itself [2]. Let us assume that we can represent the structural difference between model and process by means of an additive term. Furthermore, process measurements are always more or less corrupted by noise so that an "ideal model," which accurately and completely represents the process, will be given by

$$\dot{x}_p(\theta^*, t) = f(x_p(\theta^*, t), u(t), x_0, \theta^*, t) + e_s(\theta^*, t) \quad (3)$$

and

$$y_p(\theta^*, t) = g(x_p(\theta^*, t), u(t), x_0, \theta^*, t) + v(t) \quad (4)$$

The term $e_s(\theta^*, t)$ specifies the modeling or structural error which represents our lack of complete knowledge. It can be considered deterministic or stochastic. Measurement noise, $v(t)$, is included in equation (4) for completeness but will be assumed to be zero for the moment. The symbol θ^* represents the true parameter values of the system. Clearly, the ideal model of equations (3) and (4) can never be known exactly in practice, but it forms a reference against which actual models, like equations (1) and (2), can be judged. In this case the ideal model corresponds to the "base model" defined by Zeigler [3]. The relation between the system true model and the class of models under consideration is illustrated in Figure 1.

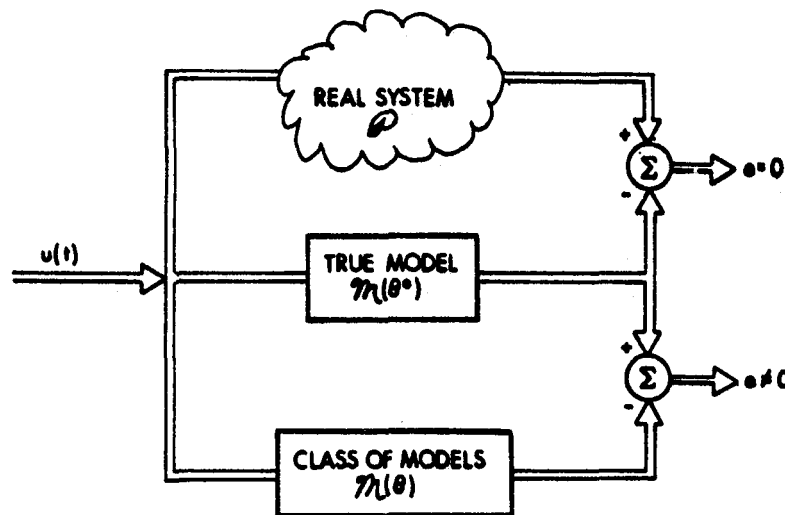


Figure 1

3. THEORETICAL IDENTIFICATION

We now define the theoretical identification problem as follows, following Zadeh [4]:

Given (a) a physical process under test, (b) a class of inputs $u(t)$, (c) a class of models $\mathcal{M}(\theta)$.

From input/output observations of ρ determine a member of \mathcal{M} which is equivalent to ρ , in the sense that its responses to all $u(t) \in \mathcal{U}$ are identical to those of ρ , i.e.,

$$y_p(t) \equiv y_m(t) \quad \forall u(t) \in \mathcal{U}, t \in [0, T] \quad (5)$$

It is evident that the solution of this identification problem is only possible if both measurement noise and structural error are identically equal to zero. If only one parameter value θ exists for which such a solution is possible, the model is set to be globally identifiable [5].

Much controversy exists in the literature concerning the theoretical identification problem. It seems to us that such problems, while interesting, have very limited usefulness in the real world.

4. PRACTICAL IDENTIFICATION

In practice, of course, the modeling error is never identically zero, and hence the model outputs can only approximate the process outputs. Let us define a scalar criterion function, $J(\mathcal{M}(\theta), \rho)$ which is a measure of the match between y_p and y_m . We can then define the real world identification problem as follows:

Definition 1. Given

- (a) a physical process under test, ρ
- (b) a class of inputs $\mathcal{U} = \{u\}$
- (c) a class of models $\mathcal{M}(\theta)$ characterized by equations (1) and (2)
- (d) a criterion function $J_p(\mathcal{M}(\theta), \rho)$
- (e) an allowable modeling error ϵ_p .

The real world identification problem consists of the determination, on the basis of input and output observations of ρ and $\mathcal{M}(\theta)$, of a model parameter vector θ for which

$$J_p(\mathcal{M}(\theta), \rho) \leq \epsilon_p$$

where ϵ_p is the allowable process identification error. If such a parameter vector can be found, we shall term the process R-W identifiable.

Note that this definition does not require exact agreement between model and process outputs. In fact, now define the notion of near equivalence by specifying a magnitude on the norm of the difference between model and process outputs. If we select a value for this norm, say δ , then we can state that:

Given that

$$\|y_p(t) - y_m(\theta, t)\| \leq \delta \quad (6)$$

the model and process are nearly equivalent. Note that criterion function appearing in Definition 1 may be the same as equation (6), or it may be an alternate criterion which measures the quality of approximation of the behavior of the process and the model. In any case, it is evident that if equation (6) is used for the criterion function in Definition 1, then we can state that models which are nearly equivalent to a given process are also R-W identifiable.

5. RELATIONS BETWEEN MODELS

In practice, we frequently approximate the model of equations (3) and (4) by a simpler and more tractable set of equations. For example, we may choose to approximate (3) by the linear model:

$$\dot{z}(\theta, t) = A(\theta)z(t) + B(\theta)u(t) + e'_s(\theta, t) \quad (7)$$

$$y_m(t) = C(\theta)z(t) + D(\theta)u(t) \quad (8)$$

where it is assumed that measurement errors are negligible. The new structural error $e'_s(\theta, t)$ in eq. (6) now includes the effects of model simplification.

Similarly, a linear high-order model may be approximated by a lower-order linear model. Let us assume that such a simpler class of models can be found, without reducing the order of the parameter vector θ , (the argument which follows can be extended to the case where the simpler model has fewer parameters than the complex model).

Consider a complex model \mathcal{M}_1 of the process which is being approximated by a simpler model \mathcal{M}_2 . Under these conditions, even in the absence of measurement noise, the models \mathcal{M}_1 and \mathcal{M}_2 will not be equivalent, since the outputs y_{m_1} of \mathcal{M}_1 will not be identical to the outputs y_{m_2} of \mathcal{M}_2 for all time in the interval of observation.

To make these ideas more precise, consider model \mathcal{M}_1 with parameter vector θ and model \mathcal{M}_2 with a different structure but the same parameter vector.

We now now define the relation between these two models independent of the quality of their approximation to the physical process.

Definition 2. Model-Model Near Equivalence (MMNE)

Two models \mathcal{M}_1 and \mathcal{M}_2 with different structures and with outputs y_{m_1} and y_{m_2} respectively are termed model-model nearly equivalent if there exists a criterion $J_m(\mathcal{M}_1(\theta), \mathcal{M}_2(\theta))$ and appropriate bounds ϵ_m and δ_m such that

$$|J_m(\mathcal{M}_1(\theta), \mathcal{M}_2(\theta))| \leq \epsilon_m \quad \text{and} \quad (9)$$

$$\|y_{m_1}(\theta, t) - y_{m_2}(\theta, t)\| \leq \delta_m \quad (10)$$

A further discussion of the near-equivalence concept and its implications is given in [6].

6. SOME SIMPLE EXAMPLES

Consider first an electrical circuit as illustrated in Fig. 2a. This is a diagram of the process. It contains a condenser with capacitance θ_1 , a coil with inductance θ_2 and a small resistor R in series. Over a given range of frequencies we assume that the current in the circuit is described by the equation (the "true" model):

$$\theta_2 I''(t) + R I'(t) + (1/\theta_1) I(t) = 0 \quad (11)$$

$$I(0) = I_0, \quad I'(0) = I'_0$$

If we neglect the small resistance of the circuit (which may represent the resistance of the coil), we obtain a model equation

$$\theta_2 i''(t) + (1/\theta_1) i(t) = 0 \quad (12)$$

$$i(0) = I_0, \quad i'(0) = I'_0$$

If one is interested in the solution only on a short time interval, the solutions of (11) and (12) may be very close. We can select a modeling error ϵ_p such that

$$|I(t) - i(t)| \leq \epsilon_p \quad (13)$$

It is evident that (13) will be satisfied only over an interval (t_0, t_f) . If, however, the final time t_f is allowed to increase, the two solutions will differ since $I(t) \rightarrow 0$ as $t \rightarrow \infty$ while $i(t)$ performs periodic oscillations with constant amplitude. The neglect of R represents the structural error and can lead to qualitatively different behavior as $t_f \rightarrow \infty$. The solution of (12) and (13) as well as the left hand side of (13) are plotted in Figure 3.

As a second example consider a linear process with time delay described by the process equation

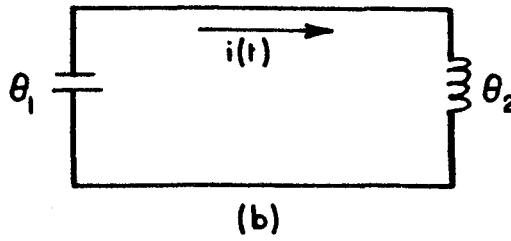
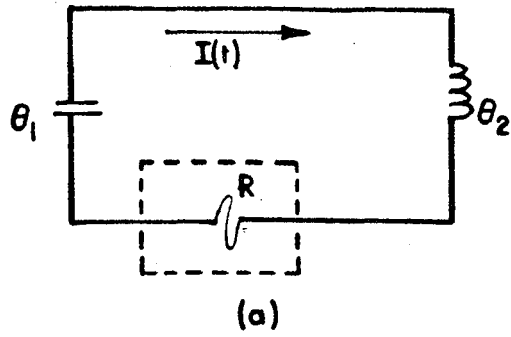


Figure 2

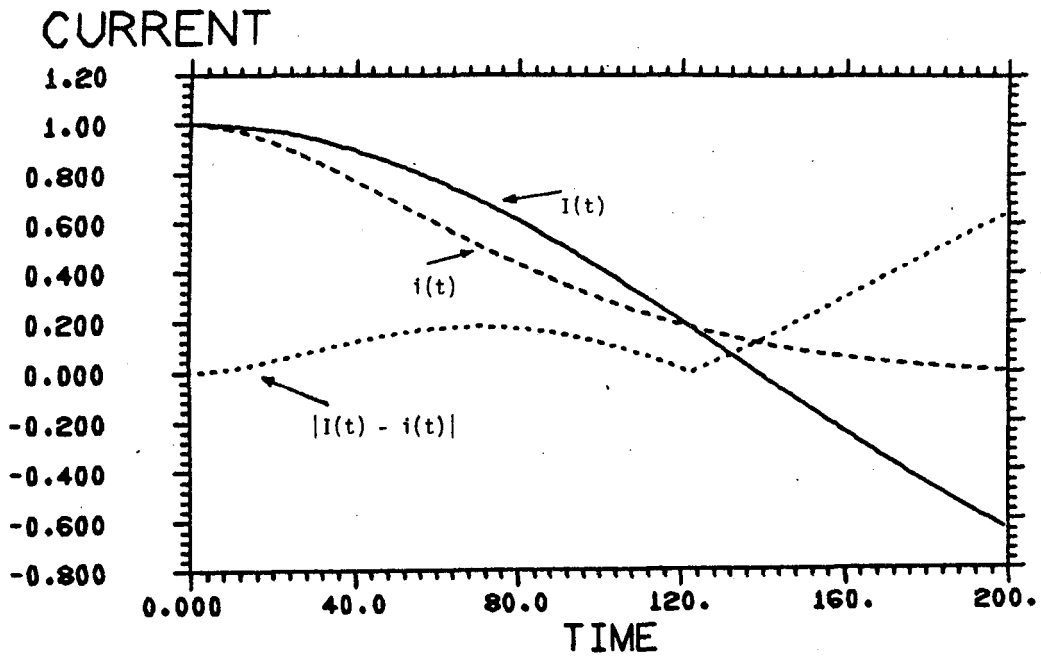


Figure 3

$$\dot{x}_p = a^* x_p + b^* u(t) + c^* x_p(t-\tau) \quad (14)$$

We assume that both c^* and τ are small and model the system as

$$\dot{x}_m = a x_m + b u(t) \quad (15)$$

where both (14) and (15) have zero initial conditions. In order to examine the effect of the structural error we compare the transfer functions of model and process. Since the time delay is small, we approximate the Laplace transform of the delay by

$$e^{-\tau s} \approx 1 - \tau s \quad (16)$$

which leads to

$$G_p(s) = \frac{X_p(s)}{U(s)} = \frac{b^* / (1 + \tau c^*)}{s - \left(\frac{a^* + c^*}{1 + \tau c^*} \right)} \quad (17)$$

while the model transfer function becomes

$$G_m(s) = \frac{X_m(s)}{U(s)} = \frac{b}{s - a} \quad (18)$$

Comparison of (17) and (18) reveals that identification of the model leads to incorrect values of the parameters a and b due to the structural error. Even if the identification is exact, the resulting values will differ from the "true" values a^* and b^* by terms which depend on the structural error parameters τ and c^* .

The above examples are very simple, but they illustrate the nature of the problem.

7. THEORETICAL RESULTS

An extensive theoretical analysis of the structural error problem has been performed for both the deterministic and stochastic case [7]. The major results in [7] can be summarized as follows:

A. Solution Error Bounds

If the time dependence of the structural error $e_s(t)$ is given, it is possible to express the solution error as a function of $e_s(t)$, i.e.,

$$|\delta y(t)| = |y_p(t) - y_m(t)| = g(e_s(t))$$

under appropriate conditions.

B. Near Equivalence

Necessary and sufficient conditions under which process and model are

near-equivalent have been found, i.e., for a given ϵ ,

$$|y_p(t) - y_m(t)| \leq \epsilon$$

C. For given values of maximum structural error $e_{s \max}$ and solution error $|\delta y|_{\max}$, bounds on the solution time have been obtained.

D. RW-Identifiability

Conditions under which the given class of models is real-world identifiable in the sense of Def. 2 have also been obtained.

The discussion of these results is beyond the scope of this paper. However, they will be published in the near future [8,9].

8. CONCLUSION

This short paper has presented a point of view on modeling and identification which includes (rather than evading) the structural difference between models and systems. We have indicated that, from this point of view, it is possible to define "near-equivalence" between process and model and to obtain meaningful theoretical results on solution error and system identification. It remains to apply these results to large realistic problems such as those involving models of complex man-machine systems.

REFERENCES

1. Eykhoff, P., "System Identification, Parameter and State Estimation," Wiley, London, 1974.
2. Bekey, G. A., "Models and Reality: Reflections on the Art and Science of Simulation," Simulation, 29; 161-164, January 1978.
3. Zeigler, B. P., Theory of Modeling and Simulation, Wiley, 1976.
4. Zadeh, L.A., "On the Identification Problem," Trans. AIEE, AC-8; 255-256, 1956.
5. DiStefano III, J.J. and Cobelli, C., "On Parameter and Structural Identifiability," IEEE Trans. Automatic Control, AC-25; 830-833, Aug. 1980.
6. Bekey, G.A. and Hadaegh, F.Y., "Equivalence Problems in Structural Identifiability," Math. and Comp. in Simulation, 24: 460-465, 1982.
7. Hadaegh, F.Y., "The Effects of Structural Error on Parameter Identifiability of Dynamical Systems," Ph.D. Dissertation, Department of Electrical Engineering, University of Southern California, June 1984.
8. Hadaegh, F.Y., and Bekey, G.A., "Structural Error and Identifiability," Proceedings of 7th IFAC Symposium on the System Identification and Parameter Estimation, 1985 (to appear).
9. Hadaegh, F.Y., and Bekey, G.A., "Near-Identifiability of Dynamical Systems," to appear in the Journal of Mathematical Biosciences.