Kybernetika

Jaromír Štěpán

The uncertainty problem in control theory. II. The internally robust procedures

Kybernetika, Vol. 26 (1990), No. 2, 122--133

Persistent URL: http://dml.cz/dmlcz/124687

THE UNCERTAINTY PROBLEM IN CONTROL THEORY
Part II. The Internally Robust Procedures

JAROMÍR ŠTĚPÁN

The problems connected with the stability of numerical processes are analyzed in the second part. The Internally Robust (IR) procedures are proposed on the basis of the concept of the theory models derived in Part I. This approach is demonstrated on the Damped Nonlinear Least Squares (DNLS) estimator derived in the literature ([16, 17]). The use of the proposed concept is illustrated by the delimitation of the applicable models of the DNLS estimator for the class of the second order systems.

6. THE INTERNALLY ROBUST PROCEDURES

As we have shown in Section 4 only the problems, which start from measured data given with the low precision, are interesting for the applications of the control theories. So we shall try to show in this section that the low accuracy of initial data may be in certain sense an advantage. But in this case one condition must be assured: It is possible to separate the data and computation precision levels. Then Hypothesis (iii) of Definition 5.2 can be fulfilled. We must distinguish in this connection at least two parts of the robustness problem. The internal robustness is connected with the stability of numerical processes. It must be first assured and then the external robustness, i.e. the robustness to errors of measured data, makes sense. In this section we shall analyze the internal robustness of control algorithms — the Internally Robust (IR) procedures.

6.1 The definition of the IR procedure

First let us introduce the intermediate result $Z((\cdot), C, d_d, d_c)$ computed for some problem from the subspace $R((\cdot), C, \sigma_L)$ with the algorithm pertinent to the theory model $T((\cdot), C, d_d, d_c)$. This intermediate result $Z((\cdot), C, d_d, d_c)$ is the numerical result of some part of the considered algorithm. If the intermediate results do not depend on the precision of input data then it must hold for the given problem and

for two different precision levels of input data, e.g. for d_d and $d_d + 1$, and for the given computation precision, e.g. $d_c = d_{cg}$, the following equality

$$Z((\cdot), C, d_d, d_{cg}) = Z((\cdot), C, d_d + 1, d_{cg}) . \tag{6.1}$$

This condition is not sufficient for the internal robustness of the CR theories because there can exist some problems for which the intermediate result cannot be computed with the computation precision d_{cg}. So the limit computation precision d_{cm} must exist for each problem from the subspace $R((\cdot), C, \sigma_L)$. The question arises how to bypass the problems with $d_{cm} > d_{cg}$. Here the backwards analysis may be useful. We are interested on the demarcation of the reliable results, i.e. on the demarcation of the subspace of results $\bar{R}((\cdot), C, \sigma_L, d_c)$ for the best possible precision of initial data $d_{dM}(q_L)$ resp. σ_{LM}. We can choose for control problems the value $d_{dM}(q_L) \doteq 2$ resp. $\sigma_{LM} \doteq 2,4 \cdot 10^{-3}$ (cf. Table 2). Now if the difference $\Delta d = d_{cg} - d_{dM}(q_L)$ is sufficiently large then it holds for $\sigma_L \geqq \sigma_{LM}$ and the given computation precision $d_c = d_{cg}$

$$\bar{R}((\cdot), C, \sigma_L, d_{cg}) \subseteq \bar{R}((\cdot), C, \sigma_{LM}, d_{cg}) \tag{6.2}$$

resp. for the given σ_{LM} and $d_{cm} > d_{cg}$

$$\bar{R}((\cdot), C, \sigma_{LM}, d_{cg}) = \bar{R}((\cdot), C, \sigma_{LM}, d_{cm}) .$$

The limiting factor for the demarcation of the subspace $\bar{R}((\cdot), C, \sigma_L, d_c)$ is here given by measured data, i.e. by the limit deviation σ_{LM}. The difference to numerical mathematics, i.e. for $d_d = d_c$ and for $d_{cm} > d_{cg}$, is given by the inclusion

$$\bar{R}((\cdot), C, d_d = d_c = d_{cg}) \subseteq \bar{R}((\cdot), C, d_d = d_c = d_{cm}) .$$

Now we can define the IR procedure:

Definition 6.1. The numerical procedure, i.e. the part of the algorithm pertinent to the model $T((\cdot), C, d_d, d_c)$ of the formal theory T_f, is internally robust with respect to the subspace of problems $R((\cdot), C, \sigma_L)$ if the following conditions are fulfilled

(i) initial data are not included in this procedure,

(ii) relation (6.1) holds for the intermediate results of all problems for the subspace $R((\cdot), C, \sigma_L)$ for the given computation precision, e.g. $d_c = d_{cg} = 16$,

(iii) the difference $\Delta d = d_c - d_{dM}(q_L)$ is so large that the delimination of the subspace $R((\cdot), C, \sigma_L)$ is given backwards by relation (6.2) for the data precision $\sigma_L \geqq \sigma_{LM}$.

Let us add that these conditions can be simply tested as we shall show later. At first sight this concept may see unrealistic. But if we take into account that the difference Δd in Hypothesis (iii) of Definition 6.1 for $d_{cg} = 16$ and $d_{dM}(q_L) = 2$ is $\Delta d = 14$ and these precision levels are realistic for the great part of practical control tasks then this concept is at least promising.

The best way, how to explain the idea of the IR procedures, is the demonstration on some identification method. Therefore we shall illustrate the proposed concept with the Damped Nonlinear Least Squares (DNLS) estimator as it was derived in the literature ([16], [17]).

6.2 The internal robustness of the DNLS estimator

First let us sketch the derivation of the Newton method to get the insight in the problems connected with the internal robustness of the DNLS estimator. The error function starting from the response of the model with the transfer function (2.3) and the measured points $\tilde{\boldsymbol{y}}$ is given by the relation

$$Q(\boldsymbol{a}) = \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{a})\boldsymbol{f}(\boldsymbol{a}), \tag{6.3}$$

where $f(k, \boldsymbol{a}) = \tilde{y}(k) - \bar{y}(k, \boldsymbol{a})$ $(k = 1, 2, \ldots, q)$.

This nonlinear error function can be approximated with the first three members of the Taylor expansion

$$Q(\boldsymbol{a} + \Delta\boldsymbol{a}) = Q(\boldsymbol{a}) + Q'(\boldsymbol{a})\,\Delta\boldsymbol{a} + \tfrac{1}{2}Q''(\boldsymbol{a})\,\Delta\boldsymbol{a}^2 + \ldots$$

The solution must be iterative and so the notation ${}^{j+1}\boldsymbol{a} = {}^{j}\boldsymbol{a} + \Delta\,{}^{j}\boldsymbol{a}$ will be used. Superscripts on the left indicate the iteration steps. Then the solution is given by the derivative of the error function

$$Q({}^{j+1}\boldsymbol{a}) = Q({}^{j}\boldsymbol{a}) + Q'({}^{j}\boldsymbol{a})\,\Delta\,{}^{j}\boldsymbol{a} + \tfrac{1}{2}Q''({}^{j}\boldsymbol{a})\,\Delta\,{}^{j}\boldsymbol{a}^2$$

at $\Delta\,{}^{j}\boldsymbol{a}$, i.e. from the condition

$$Q'({}^{j}\boldsymbol{a}) + Q''({}^{j}\boldsymbol{a})\,\Delta\,{}^{j}\boldsymbol{a} = 0 \tag{6.4}$$

we obtain the known result

$$\Delta\,{}^{j}\boldsymbol{a} = -[Q''({}^{j}\boldsymbol{a})]^{-1}\,Q'({}^{j}\boldsymbol{a}). \tag{6.5}$$

The elements of the matrices in relation (6.5) can be computed for the first partial derivatives denoted by $\partial_i\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}) = \partial\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a})/\partial\,{}^{j}a_i$ $(i = 0, 1, \ldots, \bar{n})$ resp. for the second derivatives denoted by $\partial_i\partial_e\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}) = \partial^2\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a})/\partial\,{}^{j}a_i\,\partial\,{}^{j}a_e$ $(i, e = 0.\,1, \ldots, \bar{n})$ from the relations

$$\partial_i Q({}^{j}\boldsymbol{a}) = -2\sum_{k=1}^{q}\partial_i\,{}^{j}\bar{y}(k, {}^{j}\boldsymbol{a})\,f(k, {}^{j}\boldsymbol{a}) \quad (i = 0, 1, \ldots, \bar{n}) \quad (k = 1, 2, \ldots, q) \tag{6.6}$$

resp.

$$\partial_i\partial_e Q({}^{j}\boldsymbol{a}) = 2\sum_{k=1}^{q}\partial_i\,{}^{j}\bar{y}(k, {}^{j}\boldsymbol{a})\,\partial_e\,{}^{j}\bar{y}(k, {}^{j}\boldsymbol{a}) - 2\sum_{k=1}^{q}\partial_i\partial_e\,{}^{j}\bar{y}(k, {}^{j}\boldsymbol{a})\,f(k, {}^{j}\boldsymbol{a}) \quad (i, e = 0, 1, \ldots, \bar{n}) \quad (k = 1, 2, \ldots, q). \tag{6.7}$$

The estimator based on the Newton method cannot be robust. It is shown by the analysis of the internal robustness which can be essentially simplified. Only the inversion of the Hessian matrix given by the elements according relations (6.7) must be tested. These elements are given by measured data too, i.e. $f(k, {}^{j}\boldsymbol{a})$ in the second part of relation (6.7) is given with the precision $d_d(q_L) \doteq 2$. Therefore the estimators derived from the Newton method are practically useless in spite of the fact that the pertinent formal theory is correct. Here the viewpoints of the numerical mathematics resp. the error analysis are decisive (cf. Section 3).

Similarly the main line of the computation by the great part of estimators is given by the inversion of the information matrix. Now if the elements of this matrix are given by measured data then we cannot expect that the pertinent inversion will be internally robust. The explanation is simple: we mostly lose in nontrivial cases during the computation more than two decimal places. Therefore we can formulate more general statement (cf. [16], [17]): All methods, which are based on the inversion of matrices given by measured data, cannot be in nontrivial cases internally robust, e.g. the Kalman-Bucy filter or EE estimators are not internally robust (see Paragraph 6.3).

Now if we neglect the second part of relation (6.7) then we get the Gauss-Newton method and relation (6.5) has the form

$$\Delta\, {}^{j}\boldsymbol{a} = [\boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})\, \boldsymbol{H}({}^{j}\boldsymbol{a})]^{-1}\, \boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})\, \boldsymbol{f}({}^{j}\boldsymbol{a}) \tag{6.8}$$

where $\boldsymbol{H}({}^{j}\boldsymbol{a})$ is the Jacobian matrix with elements ${}^{j}h_{ki} = {}^{j}v^{(i)}(k)$ $(k = 1, 2, \ldots, q;$ $i = 0, 1, \ldots, \bar{n})$. All estimators based on this method are internally robust. It follows from the fact that the matrix $(\boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})\, \boldsymbol{H}({}^{j}\boldsymbol{a}))$ is given with the double precision and so the pertinent inversion is numerically robust for the practically important cases (see Section 7). Let us remark that measured input signals included in sensitivity functions (2.5) do not influence the internal robustness. It follows from the fact that the Jacobian matrix can be computed with double precision for an arbitrary input signal. The pertinent resulting error must be respected in the limit deviation σ_L of the noise (cf. Paragraph 4.2 and Section 7).

Let us remark that the internal robustness of the individual iteration step does not lead automatically to the convergence of the sequence of iterations. So we must pay for the mentioned advantage, i.e. the internal robustness, by the transition to the next step. The Gauss-Newton method can be effectively used if the following condition holds (see [17])

$$(\boldsymbol{H}^{\mathrm{T}}({}^{j+1}\boldsymbol{a})\, \boldsymbol{H}({}^{j+1}\boldsymbol{a})) \approx (\boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})\, \boldsymbol{H}({}^{j}\boldsymbol{a}))\,. \tag{6.9}$$

At the present time there is no procedure how to predict $\boldsymbol{H}({}^{j+1}\boldsymbol{a})$ from the parameters of the jth iteration step. This is the main reason why the applicability of the Gauss-Newton method is so small and why different modifications are used. This problem was solved by deriving the DNLS method resp. the DNLS estimator in the literature [16] resp. [17]. Here we can conclude that the internal robustness is the necessary but not sufficient condition for the applicability of numerical procedures. The internal robustness is connected above all with the Lipschitz condition as it was formulated in Paragraph 3.2. But according to this paragraph two other requirements must be fulfilled. Firstly the solution must exist, i.e. the reliable result can be estimated from the given input data. Therefore the external robustness must be tested. Secondly the solution must be unique, i.e. the sequence $\{{}^{j}\boldsymbol{a}\}$ must converge to the global minimum. These two requirements must be analyzed to show the robustness of the DNLS estimator.

Let us set out from the resulting relations of the DNLS estimator (see [17])

$$^{j+1}\boldsymbol{a} = {}^{j}\boldsymbol{a} + {}^{j}\mu\Delta\,{}^{j}\boldsymbol{a} = {}^{j}\boldsymbol{a} + {}^{j}\mu[\boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})\,\boldsymbol{H}({}^{j}\boldsymbol{a})]^{-1}\,\boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})\,(\tilde{\boldsymbol{y}} - {}^{j}\bar{\boldsymbol{y}}) \tag{6.10}$$

and

$$^{j+1}\hat{\boldsymbol{a}} = {}^{j}\boldsymbol{a} - {}^{j}\mu\Delta\,{}^{j}\boldsymbol{a} = {}^{j}\boldsymbol{a} + {}^{j}\mu[\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\,\boldsymbol{H}_L({}^{j}\boldsymbol{a})]^{-1}\,\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\,(\tilde{\boldsymbol{y}} - {}^{j}\bar{\boldsymbol{y}})\,, \tag{6.11}$$

where $\mu \in (0, 1)$ is a damping factor. The Jacobian matrix $\boldsymbol{H}_L({}^{j}\boldsymbol{a}) = -\boldsymbol{H}({}^{j}\boldsymbol{a})$ corresponds to the linear case, i.e. to the case with known sensitivity functions ${}^{j}v^{(i)}(t)$ $(i = 0, 1, \ldots, \bar{n})$ in relation (2.6). This linear case – this linear Gauss estimator – allows to demarcate the region in which the linearization of the pertinent nonlinear function can be used and so the right damping factor ${}^{j}\mu$ can be calculated.

To be able to test the external robustness we must know that some simple method if possible linear can be used. So the following proposition will be useful:

Proposition 6.1. The DNLS estimator is quasi-linear, i.e. the sequence $\{{}^{j}\boldsymbol{a}\}$ of the estimates is governed near the global minimum by the linear solution given by relation (6.11).

Proof. We start from the functions which are given by relations (6.10) and (6.11) multiplied from the left with the matrix $\boldsymbol{H}_L({}^{j}\boldsymbol{a})$

$$^{j}\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}, {}^{j+1}\boldsymbol{a}, {}^{j}\mu) = \boldsymbol{H}_L({}^{j}\boldsymbol{a})\,{}^{j+1}\boldsymbol{a} = {}^{j}\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}) - {}^{j}\mu\Delta\,{}^{j}\bar{\boldsymbol{y}} \tag{6.12}$$

and

$$^{j}\boldsymbol{z}({}^{j}\boldsymbol{a}, {}^{j+1}\hat{\boldsymbol{a}}, {}^{j}\mu) = \boldsymbol{H}_L({}^{j}\boldsymbol{a})\,{}^{j+1}\hat{\boldsymbol{a}} = {}^{j}\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}) + {}^{j}\mu\Delta\,{}^{j}\bar{\boldsymbol{y}}\,, \tag{6.13}$$

where ${}^{j}\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}) = \boldsymbol{H}_L({}^{j}\boldsymbol{a})\,{}^{j}\boldsymbol{a}$, $\Delta\,{}^{j}\bar{\boldsymbol{y}} = \boldsymbol{H}_L({}^{j}\boldsymbol{a})\,\Delta\,{}^{j}\boldsymbol{a}_L$ and $\Delta\,{}^{j}\boldsymbol{a}_L = -\Delta\,{}^{j}\boldsymbol{a} = {}^{j}\boldsymbol{a} - {}^{j+1}\boldsymbol{a}$. The linear solution is given by relation (6.13). It follows from relation (6.11) which can be written for ${}^{j}\mu = 1$ and with respect to the relation

$$^{j}\boldsymbol{a} = [\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\,\boldsymbol{H}_L({}^{j}\boldsymbol{a})]^{-1}\,\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\,{}^{j}\bar{\boldsymbol{y}} \tag{6.14}$$

in the known form

$$^{j+1}\hat{\boldsymbol{a}} = [\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\,\boldsymbol{H}_L({}^{j}\boldsymbol{a})]^{-1}\,\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\,\tilde{\boldsymbol{y}}\,. \tag{6.15}$$

If we are sufficiently near to the linear case then it holds (see [17])

$$\begin{aligned} {}^{j+1}\bar{\boldsymbol{y}}({}^{j+1}\boldsymbol{a}, {}^{j}\mu) &= \boldsymbol{H}_L({}^{j+1}\boldsymbol{a})\,{}^{j+1}\boldsymbol{a} \doteq {}^{j}\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}, {}^{j+1}\boldsymbol{a}, {}^{j}\mu) + 2\,{}^{j}\mu\Delta\,{}^{j}\bar{\boldsymbol{y}} = \\ &= \boldsymbol{H}_L({}^{j}\boldsymbol{a})\,{}^{j+1}\boldsymbol{a} + 2\,{}^{j}\mu\,\boldsymbol{H}_L({}^{j}\boldsymbol{a})\,\Delta\,{}^{j}\boldsymbol{a}_L\,, \end{aligned} \tag{6.16}$$

where the change $\boldsymbol{H}_L({}^{j}\boldsymbol{a})$ into $\boldsymbol{H}_L({}^{j+1}\boldsymbol{a})$ is approximately given with the term $2\,{}^{j}\mu\Delta\,{}^{j}\bar{\boldsymbol{y}}$. Now the nonlinear solution in the jth iteration step is given by the linear etalon, i.e. it holds with respect to relations (6.12), (6.13) and (6.16)

$$^{j+1}\bar{\boldsymbol{y}}({}^{j+1}\boldsymbol{a}, {}^{j}\mu) \doteq {}^{j}\bar{\boldsymbol{y}}({}^{j}\boldsymbol{a}) + {}^{j}\mu\Delta\,{}^{j}\bar{\boldsymbol{y}} = {}^{j}\boldsymbol{z}({}^{j}\boldsymbol{a}, {}^{j+1}\hat{\boldsymbol{a}}, {}^{j}\mu)\,. \tag{6.17}$$

If the difference $\Delta\,{}^{j}\bar{\boldsymbol{y}}$ is small then the linear etalon ${}^{j}\boldsymbol{z}({}^{j}\boldsymbol{a}, {}^{j+1}\hat{\boldsymbol{a}}, {}^{j}\mu)$ well approximates the resulting nonlinear function ${}^{j+1}\bar{\boldsymbol{y}}({}^{j+1}\boldsymbol{a}, {}^{j}\mu)$. □

The sequence $\{{}^{j}\boldsymbol{a}\}$ is governed in the endsteps, i.e. for $\Delta\,{}^{j}\boldsymbol{a} \Rightarrow 0$, by the linear case. From this fact the important conclusion can be drawn: The linear estimation theory

can be used by testing the external robustness in all situations in which ${}^{j}\boldsymbol{a}$ is sufficiently near to the global minimum.

Now let us prove the uniqueness of the solution. The asterisk on the left indicates the coefficients and functions pertinent to the global minimum.

Proposition 6.2. The DNLS estimator has the global minimum $*\boldsymbol{a}$ for $\Delta\, {}^{j}\boldsymbol{a} = 0$, i.e. for $j \to \infty$ we obtain $*\bar{\boldsymbol{y}} = *\boldsymbol{z}$. Local minima cannot exist.

Proof. Let us sketch the proof according to the literature [17]. We shall start from the linear regression function ${}^{j}\boldsymbol{z} \in \bar{\bar{R}}({}^{j}\bar{\boldsymbol{y}}, {}^{j}\boldsymbol{z}, \sigma_L, d_c)$ and the fact that the spaces L_p are for $1 < p < \infty$ strict convex. Then only the unique best solution ${}^{j}\boldsymbol{z} \in \bar{\bar{R}}({}^{j}\tilde{\boldsymbol{y}}, {}^{j}\boldsymbol{z}, \sigma_L, d_c)$ can exist for the measured points $\tilde{\boldsymbol{y}}$, i.e. according to relation (5.1) for $\boldsymbol{y}(\sigma_L) \in \in R(\boldsymbol{y}, \sigma_L)\,(\bar{R} \subset R \subset l_2)$. Now it holds with respect to Proposition 6.1 $*\bar{\boldsymbol{y}} = *\boldsymbol{z}$ for $j \Rightarrow \infty$. So only one minimum can exist. □

The DNLS estimator associates the numerical robustness with the advantages of the linear theory. The details connected with the convergence resp. the practical use of this estimator can be found in the literature ([16], [17]).

6.3 Robust models of identification theories

Let us complete the discussion of the robust theory models from Section 5. It follows from the previous discussion that only the demarcated subspaces $R((\cdot), C, \sigma_L)$ of problems with reliable results given by the subspace $\bar{R}((\cdot), C, \sigma_L, d_c)$ can be practically used, and this demarcation must be ascertained by the experimentation. Therefore testing algorithms resp. procedures with respect to uncertainties is an unseparable part of the error analysis. First let us demonstrate this testing on a simple case – on the quadratic equation from Paragraph 3.1.

The numerical procedure for the computation of the root x_2 from equation (3.1) is based on the theory – on the relation

$$x_2 = -b + \sqrt{(b^2 - c)}\,. \tag{6.18}$$

So the models of this "theory" $T((\text{comp. } \bar{x}_2), d_d, d_c)$ generate the results with the different errors according to Table 1. One can obtain for different d_d and d_c results with different reliability, e.g. $\Theta_k < 100\%$ for $\alpha < 7$, or decide that the problem is unsolvable on the given precision levels, e.g. $\Theta_k \geqq 100\%$ for $\alpha \geqq 7$ (cf. Table 1) by the computation according to the theory model $T((\text{comp. } \bar{x}_2), d_d = d_c = 13)$.

The error analysis of the estimators is more complicated. Here we must solve this problem at least in two steps. First the internal robustness must be assured. So we have shown in Paragraph 6.2 that the main computation line of the DNLS estimator is the IR procedure (cf. Definition 6.1), i.e. the matrix inversion $[\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\, \boldsymbol{H}_L({}^{j}\boldsymbol{a})]^{-1}$ resp. the term $[\boldsymbol{H}_L^{\mathrm{T}}({}^{j}\boldsymbol{a})\, \boldsymbol{H}_L({}^{j}\boldsymbol{a})]^{-1}\, \boldsymbol{H}^{\mathrm{T}}({}^{j}\boldsymbol{a})$ is the IR procedure. Therefore relation (6.1) must hold for this intermediate result, e.g. for the given computation precision

$d_c = d_{cg} = 16$ and for the arbitrary data precision $d_{dM}(q_L) < 3$ resp. $\sigma_L > \sigma_{LM} = 2{\cdot}4 \,.\, 10^{-4}$

$$Z((\text{comp.}\,[\boldsymbol{H}_L^{\mathrm{T}}({}^j\boldsymbol{a})\,\boldsymbol{H}_L({}^j\boldsymbol{a})]^{-1}), C, \sigma_L, d_{cg}) = \text{const.} \tag{6.19}$$

So the internal robustness of the main computation line can be very simply tested (see Section 7).

Now let us show the opposite situation. The difference equation of the linear Equation Error (EE) methods (cf. [5], [17], [19]) is given by the relation

$$\tilde{y}(k) = \boldsymbol{\theta}^{\mathrm{T}}\,\boldsymbol{r}(k) + w(k)\,, \tag{6.20}$$

where $\boldsymbol{\theta} = [\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_m]^{\mathrm{T}}$ and

$$\boldsymbol{r}(k) = [-\tilde{y}(k-1, \ldots, -\tilde{y}(k-n), u(k-1), \ldots, u(k-m)]^{\mathrm{T}}\,.$$

The parameter vector $\boldsymbol{\theta}$ is to be estimated from the measured points $\tilde{y}(k)$ and $r(k)$. The "equation error" is then given by

$$Q_E(\boldsymbol{\theta}) = \frac{1}{q}\sum_{k=1}^{q}[\tilde{y}(k) - \boldsymbol{\theta}^{\mathrm{T}}\,\boldsymbol{r}(k)]^2\,. \tag{6.21}$$

The numerical solution is given by the relation

$$\boldsymbol{\theta}(q) = [\sum_{k=1}^{q}\boldsymbol{r}(k)\,\boldsymbol{r}^{\mathrm{T}}(k)]^{-1}\sum_{k=1}^{q}\boldsymbol{r}(k)\,\tilde{y}(k)\,. \tag{6.22}$$

The elements of the matrices on the right side of relation (6.22) are the measured data given with the precision $d_d(q_L) \doteq 2$, i.e. $\sigma_L \doteq 2{\cdot}4 \,.\, 10^{-3}$. Therefore the main computation line, i.e. the computation of the inverse $[\sum_{k=1}^{q}\boldsymbol{r}(k)\,\boldsymbol{r}^{\mathrm{T}}(k)]^{-1}$, cannot be in nontrivial cases the IR procedure.

We have discussed till now only the internal robustness of estimators. But the ultimate demarcation of solvable problems by all theories is connected with the external robustness (cf. relation (5.2)). So the following definition of the robust theory models will be useful:

Definition 6.2. The model of the theory is robust with respect to the subspace of results $\bar{R}((\cdot), C, \sigma_L, d_c)$ if the following conditions are fulfilled

(i) the theory model was derived from the CR theory,

(ii) the main computation line is the IR procedure (cf. Definition 6.1),

(iii) the solution exists on the precision level of initial data in some sense, e.g. in sense of the null-hypothesis (cf. relation (5.4)).

The external robustness of the DNLS estimator from Paragraph 6.2 is given only by the error difference $\tilde{\boldsymbol{y}} - {}^j\bar{\boldsymbol{y}}$ in relation (6.11). If we assume the normal noise

(cf. Paragraph 4.3) then we can compute the standard deviations of the estimates $a_i\,(i = 0, 1, \ldots, \bar{n})$ in a simple way. The pertinent variances are given by the diagonal elements ${}^j\gamma^{ss}\,(s = i + 1)$ of the inverse $\left[\boldsymbol{H}_L^{\mathrm{T}}({}^j\boldsymbol{a})\,\boldsymbol{H}_L({}^j\boldsymbol{a})\right]^{-1}$

$$D({}^j a_i) = \sigma^2\,{}^j\gamma^{ss}\,, \tag{6.23}$$

where the estimate of the data mean error resp. the limit standard deviation σ_L must be approximatelly known. We see that the external robustness is here closely connected with the internal robustness. We obtain with respect to relation (6.19)

$$D({}^j a_i)/\sigma^2 = {}^j\gamma^{ss} = \text{const.} \tag{6.24}$$

The practical situations are more complicated (cf. relation (4.1)). Let us remark in this connection that the considered approach is mostly sufficient to get the insight in the considered problem as we shall show in the next section.

Let us emphasize that the reliability of the derived end-results is the matter of agreement. E.g. the identified model of the system must be appropriate for the solution of the subsequent synthesis problem (see [15]). The discussion of this topic is beyond the scope of this paper.

7. THE APPLICABLE MODELS OF THE DNLS ESTIMATOR

Let us illustrate the proposed concept of the theory models with some examples. Let us show how it is possible to demarcate the regions in which the DNLS estimator models $T(\text{(DNLS estim.)}, C, \sigma_L, d_c)$ can be used. We shall analyze the class of the second order systems of the type (2.3) $R_\varkappa(\boldsymbol{y}(\varkappa),\ \varkappa \in (0; 0{\cdot}25))$ $(R_\varkappa \subset R)$, i.e. with real roots $s_i\,(i = 1, 2)$, given by the transfer function

$$F_\varkappa(s) = 1/(1 + 10s + 100\varkappa s^2)\,. \tag{7.1}$$

The subspace of pertinent results $\bar{R}(\bar{\boldsymbol{y}}(\boldsymbol{a}, \varkappa), C, \sigma_L, d_c)$, i.e. the models of the considered system identified from the measured points $\tilde{\boldsymbol{y}} \in \bar{R}$ resp. $\boldsymbol{y}(\varkappa, \sigma_L) \in R(\boldsymbol{y}(\varkappa), C, \sigma_L)$ (cf. relation (5.1)), is given by the class of the transfer functions

$$\bar{F}_\varkappa(s) = 1/(a_0 + a_1 s + \varkappa a_1^2 s^2)\,. \tag{7.2}$$

where $\varkappa = a_2/a_1^2$. So the coefficient vector $\boldsymbol{a}$ should be estimated according to relation (5.2).

Table 3 contains important parameters computed with the double precision PL 1 program based on the algorithm from the literature [16] by the following conditions $C_\varkappa$:

The error of measured data was simulated by a round-off noise, i.e. the points of the unit impulse responses $h(k)\,(k = 1, 2, \ldots, q)$ for $q = 25$, $\Delta\tau = 0{\cdot}8$ and $\Delta t = 20$, i.e. for the time scale $t = [0{\cdot}4, 1{\cdot}2, \ldots, 19{\cdot}6]$, were given on d_d decimal places $(d_d = 1, 2, 3)$. So we consider with respect to Paragraph 4.2 for $q_L = 25$ the limit

Table 3.

$\varkappa$	d_d	$Z({}^1a_2)$	$\bar{\sigma}$	$\sqrt{D({}^1a_2)}$	${}^1s_2$	$\sqrt{D(*a_2)}$	$\Theta(*a_2)$
	3		$3{\cdot}0 \cdot 10^{-4}$	$6{\cdot}0 \cdot 10^{-2}$	$6{\cdot}0 \cdot 10^{-3}$	$6{\cdot}1 \cdot 10^{-2}$	0·61
10^{-1}	2	199·9	$2{\cdot}8 \cdot 10^{-3}$	$5{\cdot}6 \cdot 10^{-1}$	$5{\cdot}6 \cdot 10^{-2}$	$6{\cdot}1 \cdot 10^{-1}$	6·1
	1		$3{\cdot}1 \cdot 10^{-2}$	1·8	$1{\cdot}8 \cdot 10^{-1}$	1·9	19
	3		$2{\cdot}5 \cdot 10^{-4}$	$2{\cdot}2 \cdot 10^{-2}$	$2{\cdot}2 \cdot 10^{-2}$	$2{\cdot}2 \cdot 10^{-2}$	2·2
10^{-2}	2	88·2	$2{\cdot}9 \cdot 10^{-3}$	$2{\cdot}6 \cdot 10^{-1}$	$2{\cdot}6 \cdot 10^{-1}$	$2{\cdot}6 \cdot 10^{-1}$	26
	1		$3{\cdot}1 \cdot 10^{-2}$	2·7	2·7	1·5	150
	3		$2{\cdot}7 \cdot 10^{-4}$	$2{\cdot}3 \cdot 10^{-2}$	$2{\cdot}3 \cdot 10^{-1}$	$2{\cdot}3 \cdot 10^{-2}$	23
10^{-3}	2	83·4	$2{\cdot}9 \cdot 10^{-3}$	$2{\cdot}5 \cdot 10^{-1}$	2·5	div.	div.
	1		$3{\cdot}2 \cdot 10^{-2}$	2·6	26	1·5	1500

precision

$$\sigma_L^2 = \sigma^2 q_L = 10^{-2d}/(2\pi e) \cdot q_L \,. \tag{7.3}$$

In this way the identification of the considered systems is tested for the use of the measurement instruments of the accuracy classes 1%, 0·1% and 0·01% (cf. Table 2).

The starting transfer function was chosen in the form

$${}^0\bar{F}_\varkappa(s) = 1/(1{\cdot}01 + 10s + 100\varkappa s^2) \,. \tag{7.4}$$

The intermediate results $Z({}^1a_2) = {}^1\gamma^{33}$ in Table 3 corroborate the internal robustness of the DNLS estimator with respect to relation (6.1). The mean deviation $\sqrt{D({}^1a_2)}$ and the pertinent relative standard deviation ${}^1s_2$ show the influence of the different precisions of the input data for the same $Z({}^1a_2)$.

The standard deviation σ of the round-off noise can be here computed directly according to the relation

$$\bar{\sigma}^2 = 1/q \sum_{k=1}^{q} [\tilde{h}(k) - h(k)]^2 \,. \tag{7.5}$$

So the standard deviation σ calculated according to relation (4.3) gives here the lower boundary.

The end-results are characterized by the mean deviation $\sqrt{D(*a_2)}$. The pertinent error was calculated according to the relation

$$\Theta(*a_2) = 1/\varkappa \sqrt{D(*a_2)} \quad [\%] \,. \tag{7.6}$$

These values start from the results of the jth iteration step with ${}^ja_2 \approx *a_2$, i.e. for $\Delta\, {}^ja_2 < 10^{-3}$, and from the estimate of the standard deviation of the noise given by the known relation

$$\sigma^2 = \frac{1}{q - \bar{n} - 1} \sum_{k=1}^{q} [\tilde{h}(k) - {}^j\bar{h}(k)]^2 \,. \tag{7.7}$$

We can summarize the results from Table 3 with the following proposition:

Proposition 7.1. The applicable models of the DNLS estimator T((DNLS estim.), $C_\varkappa$, σ_L, d_{cg}) are given for the class of systems $R_\varkappa$, for the conditions $C_\varkappa$ and for three limit deviations $\sigma_L = 4{\cdot}8 \,.\, 10^{-p}$ $(p = 3, 4, 5)$ of measured data with the following subspaces of problems

$$R_{\varkappa p} = \{(h(\varkappa, \sigma_L): \varkappa \in (4{\cdot}04 \,.\, 10^{(1-p)}; 0{\cdot}25), \sigma_L = 4{\cdot}8 \,.\, 10^{-p}, d_{cg}\} \,. \tag{7.8}$$

Proof. Firstly we prove according to Definition 6.2 that the DNLS estimator is the CR theory (cf. Definition 5.2), i.e. the pertinent formal theory is consistent. The validity of Hypothesis (i) resp. (ii) of Definition 5.2 follows from Proposition 6.2. If the main line of computation is the IR procedure according to Definition 6.1 then Hypothesis (iii) of Definition 5.2 is fulfilled. Now Hypothesis (i) resp. (ii) of Definition 6.1 are corroborated with the intermediate results $Z({}^1a_2)$ according to relation (6.19) (cf. Table 3).

Then we must test the limit deviation σ_{LM} with respect to the limit computation precision d_{cm} according to Hypothesis (iii) of Definition 6.1. The simple test shows that the case with the parameter $\varkappa_M = 10^{-4}$ must be computed with the computation precision $d_c > d_{cg}$. Here the critical procedure is the computation of the exponentials to ${}^j\bar{h}(k, \boldsymbol{a})$ resp ${}^jv^{(i)}(k, \boldsymbol{a})$ $(k = 1, 2, \ldots, q)$ (cf. [6]).

The last step is given by testing Hypothesis (iii) of Definition (6.2). Here we must tests the limit identifiability of the coefficient estimates a_i $(i = 0, 1, 2)$ (see [12], [13]). It is mostly sufficient to test the last coefficient estimate a_2. The pertinent element ${}^*\gamma^{33} = 7, 1 \,.\, 10^3$ of the inverse $[\boldsymbol{H}_L^{\mathrm{T}}({}^j\boldsymbol{a})\, \boldsymbol{H}_L({}^j\boldsymbol{a})]^{-1}$ is practically constant for all considered cases. So we obtain for the condition ${}^*s_2 \leqq 0{\cdot}5$ and for the normed standard limit deviations $\sigma_L = 4{\cdot}8 \,.\, 10^{-p}$ the boundaries $\varkappa = 4{\cdot}04 \,.\, 10^{(1-p)}$ in relation (7.9). Let us add that these boundaries fulfil Hypothesis (iii) of Definition 6.1, i.e. $\varkappa = 4{\cdot}04 \,.\, 10^{(1-p)} > \varkappa_M = 1 \,.\, 10^{-4}$ for $p \leqq 5$. □

If we repeat this test for the step responses then we obtain the insight on the applicability of the class of considered systems in the practically important region of input signals.

This test can be simply extended on the high order systems $(\bar{n} > 2)$. The applicability region resp. the validity of the DNLS estimator model is here mostly given with the limit identifiability of the last coefficient estimate $a_{\bar{n}}$ (see [11], [12]).

8. CONCLUSION

In this paper it was shown that the present crisis of the control theory is closely connected with the incorporation of an uncertainty band in formal theories. Two kinds of the uncertainties or better two precision levels must be here distinguished. The stability of numerical processes is mostly connected with the simple resp. double precision of the computation (cf. Paragraph 4.2 and Table 1). The low precision of

measured data, i.e. $d_d(q_L) \doteq 2$, is given with the low accuracy of the operational measurements (cf. Table 2). The analysis of these problems in Sections 3 and 4 has shown that the solution of control problems can be successful if we are able to separate these two so different precision levels. Therefore the concept of the theory models (cf. Definition 5.1 and 5.2) and the internally robust procedures (cf. Definition 6.1) was introduced in Section 5 and 6. The realizability of this concept was demonstrated on the DNLS estimator (see Paragraph 6.2) and by the examples in Section 7.

Let us add two remarks:

(i) The concept of the theory models can simply explain among others the success and the limitations of the classical simple controllers, i.e. the P, PI resp. PID controllers, in practical tasks. Here we must mostly control according to the small signal deviations, i.e. near the zero precision (cf. Table 2), and so the mathematical theory is useless in accordance with the analysis in Section 5. The pertinent subspace of results is empty (cf. relation (5.2)), i.e. $R((\cdot), C, \sigma_L, d_c) = \emptyset$. The control is here robust because the signal is put down in the uncertainty band (cf. [11], [13]).

(ii) The second remark concerns with the use of the artificial intelligence by the solution of control problems. The use of the artificial intelligence allows here to overcome the troubles with the uncertainties in sense of the theory models. The formal theories cannot generate the facts – they can generate only the forecasts and these forecasts can be verified by the use of the artificial intelligence (cf. [13]). Roughly speaking it holds in the modern control theory as a rule: one system – one strategy – one controller. On the other side the concept of the theory models uses the following scheme: one system – different strategies – different theory models according to the different amount of information in measured signals – the greater number of different controllers (cf. [13]).

(Received May 4, 1989.)

REFERENCES

[1] M. Athans: Advances and open problems on the control of large scale systems. In: Proc. of VII IFAC Congress (Helsinki 1978), Pergamon Press, Oxford 1978, pp. 2371–2382.

[2] G. C. Goodwin: Some observations on robust estimation and control. In: Proc. of 7th IFAC/IFORS Symp. on Identification and Parameter Estimation, York 1985, Pergamon Press, Oxford 1985.

[3] P. V. Kokotovic: Control theory in the 80's: Trends in feedback design. Prepr. IX IFAC Congress, Vol. VI, Budapest 1984.

[4] S. Körner: Experience and Theory. Routledge & Kegan, London 1966.

[5] L. Ljung and T. Söderström: Theory and Practice of Recursive Identification. MIT Press, Massachusetts 1983.

[6] C. Moler and Ch. van Loan: Nineteen dubious ways to compute the exponential of a matrix. SIAM Rev. *20* (1978), 801–836.

[7] K. R. Popper: The Logic of Scientific Discovery. Hutchinson, London 1968.

[8] A. Ralston: A First Course in Numerical Analysis. McGraw-Hill, New York 1965.

[9] J. R. Rice: Numerical Methods, Software, and Analysis. McGraw-Hill, Tokyo 1983.

[10] J. Štěpán: Some problems of systems identification. Kybernetika *7* (1971), 2, 133–155.

[11] J. Štěpán: Messgenauigkeit und Regelungstheorie. Regelungstechnik und Prozess-Datenverarbeitung *20* (1972), 203—207.
[12] J. Štěpán: Uncertainty problem in cybernetics (in Czech). In: Repr. of the Conference on Cybernetics, Prague, November 1976.
[13] J. Štěpán: Control of regulator-configurations in large scale systems. In: Fourth Formator Symposium on Mathematical Methods for the Analysis of Large Scale Systems (J. Beneš, L. Bakule, eds.), Academia, Prague 1983.
[14] J. Štěpán: Nonlinear approximation in control problems. Kybernetika *19* (1983), 6, 491—504.
[15] J. Štěpán: Approximation versus identification. In: Proc. of 7th IFAC/IFORS Symp. on Identification and Parameter Estimation, York 1985, Pergamon Press, Oxford 1985.
[16] J. Štěpán: A new method for the nonlinear approximation of signals. Part I: The optimal damping factor. Kybernetika *22* (1986), 5, 425—438. Part II: The convergence problem. Kybernetika *22* (1986), 6, 503—513.
[17] J. Štěpán: Robust quasi-linear system identification. Kybernetika *24* (1988), 4, 259—277.
[18] V. N. Tutubalin: Theory of Probability (in Russian). Mir, Moscow 1972.
[19] P. Young: Recursive Estimation and Time-Series Analysis. Springer-Verlag, Berlin—Heidelberg—New York 1984.
[20] J. Zaborszky: Development of systems science — past, present, and future. Prepr. IX. IFAC Congress, Vol. VI. Budapest 1984.

Ing. Jaromír Štěpán, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.