KYBERNETIKA — VOLUME 27 (1991), NUMBER 5

ASSUMPTION EVALUATION IN APPROXIMATE MODELLING[1,2]

SERGIO BEGHELLI, ROBERTO P. GUIDORZI, UMBERTO SOVERINI

System identification can be considered as an approximate modelling problem on the basis of observed data and *a priori* knowledge. Assumptions are often introduced on the set of the considered models and on the characteristics of the noise affecting the data; the validity of such assumptions must then be evaluated on the basis of the identified model family. This paper refers to the Frisch scheme applied to dynamic system identification and introduces a procedure to evaluate the violation of the assumptions underlying the considered scheme.

1. INTRODUCTION

Identification is the process of constructing mathematical models of dynamic systems on the basis of measured data.

It is, in general impossible to obtain an exact mathematical description of the underlying system, because of its complexity and/or the incompleteness of the available data; in this sense it is, therefore, more reasonable to consider identification as a problem of approximate modelling.

Many traditional identification techniques justify the deviations between the identified model and the data by means of statistical considerations, and the model is often validated on the basis of assumptions which are *a priori* unverifiable and certainly not detectable on the basis of the available data.

Recently, Willems [1] has developed an alternative approach based on the concepts of model complexity and misfit between a model and the data, avoiding any statistical consideration. In his work he introduced a misfit function, based on equation error, and an approximate modelling procedure mainly oriented to description purposes.

The identification technique here described leads to descriptive models of dynamic

[1] This research has been supported by the Ministry for University and Scientific and Technological Research, Rome, under project Model Identification, Control Systems and Signal Processing.

[2] Presented at the IFAC Workshop on System Structure and Control held in Prague during 25—27 September 1989.

systems by extending the scheme originally suggested by Frisch [2] and reproposed by Kalman [3] in the algebraic case. The result of this approach is a family of admissible models which can explain the data within the assumptions on the noise. Following the same scheme other authors, e.g. Deistler [4], used Errors-in-Variables models for identification of dynamic systems, but the approximate modelling aspect seems not to be incorporated in their formulation.

This paper introduces a consistent procedure to mathematically describe the validity of the assumptions introduced in approximate modelling. The approach is here developed for single-input single-output models, but similar results can be obtained in the multivariable case.

2. DYNAMIC SYSTEM IDENTIFICATION

The considered models are of the type

$$\hat{y}(t+n) = \sum_{i=0}^{n-1} \alpha_i \hat{y}(t+i) + \sum_{i=0}^{n} \beta_i \hat{u}(t+i) \tag{2.1}$$

where $\hat{y}(\cdot)$ and $\hat{u}(\cdot)$ are the output and the input of the system respectively, n the order, α_i and β_i its parameters.

Let us assume that the system generating the data belongs to the considered model set. In absence of noise on the data, when a finite sequence of the input-output variables observed N times with a constant sampling interval is available, the order and the parameters of the system can be computed as follows.

Let us define the matrix $\hat{X}_k$ as

$$\hat{X}_k = \begin{bmatrix} \hat{y}(0) & \dots & \hat{y}(k) & \hat{u}(0) & \dots & \hat{u}(k) \\ \hat{y}(1) & \dots & \hat{y}(k+1) & \hat{u}(1) & \dots & \hat{u}(k+1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \hat{y}(L) & \dots & \hat{y}(k+L) & \hat{u}(L) & \dots & \hat{u}(k+L) \end{bmatrix} \tag{2.2}$$

where $k + L \leqq N - 1$ and $L \geqq 2k + 1$. The time-invariance of the system allows the assumption that the initial time is equal to zero.

Let us define matrix $\hat{\Sigma}_k$ partitioned in the following way

$$\hat{\Sigma}_k = \frac{1}{L+1} \hat{X}_k^{\mathrm{T}} \hat{X}_k = \frac{1}{L+1} \begin{bmatrix} \hat{\Sigma}_k(\hat{y}\hat{y}) & \hat{\Sigma}_k(\hat{y}\hat{u}) \\ \hat{\Sigma}_k(\hat{u}\hat{y}) & \hat{\Sigma}_k(\hat{u}\hat{u}) \end{bmatrix}. \tag{2.3}$$

On the assumption that the vectors of matrix (2.2) obtained from input samples are linearly independent, the order n of the system is equal to the minimum value of the integer k corresponding to a singular (non-negative definite) matrix $\hat{\Sigma}_k$. Consequently, the parameter vector $p = [\alpha_0, \dots, \alpha_{n-1}, 1, \beta_0, \dots, \beta_n]^{\mathrm{T}}$ is univocally determined by the relation $\hat{\Sigma}_n p = 0$, i.e. p belongs to the null space (kernel) of $\hat{\Sigma}_n$.

In presence of noise the previous procedure would obviously be useless since matrices $\hat{\Sigma}_k$ would always be non-singular. To solve the problem, some assumptions

on the characteristics of the noise affecting the data must be introduced. It is, in general, very difficult (if possible at all) to establish the validity of such assumptions on the basis of the data available and this has led Kalman to call them, in the algebraical case, *prejudices*. Among the several procedures proposed in the literature, we will deal, in the following, with the scheme originally suggested by Frisch [2] in the algebraic case, reproposed by Kalman [3] and later extended to dynamic system identification [5].

The main feature of this approach regards the determination of models from input-output sequences without introducing any statistical assumption on the data.

The scheme is based on the following hypotheses:

– input and output variables are affected by additive noise

$$u(t) = \hat{u}(t) + \tilde{u}(t)\,, \quad y(t) = \hat{y}(t) + \tilde{y}(t)\,; \tag{2.4}$$

– for every k the generic positive definite matrix Σ_k associated with the input-output noise-corrupted sequences can be decomposed as

$$\Sigma_k = \hat{\Sigma}_k + \tilde{\Sigma}_k \tag{2.5}$$

where

$$\tilde{\Sigma}_k = \operatorname{diag}\left[\tilde{\sigma}_y I_{k+1},\, \tilde{\sigma}_u I_{k+1}\right] \geqq 0\,. \tag{2.6}$$

From an algebraic point of view, conditions (2.5) and (2.6) are equivalent to assuming that the columns of matrix $\tilde{X}_k$, whose entries are the samples $\tilde{y}(\cdot)$ and $\tilde{u}(\cdot)$ arranged as in (2.2), are orthogonal to each other and orthogonal also to the vectors of matrix $\hat{X}_k$.

From decomposition (2.5) of Σ_k it follows that the solution of the identification problem can be obtained in two separate steps:

– determination of the noise model, described by the diagonal terms of $\tilde{\Sigma}_k$;
– determination of the process model, from the algebraic links among the vectors of the non-negative definite matrix $\hat{\Sigma}_k$.

Note that when system (2.1) is stable and $L \to \infty$, conditions (2.5) and (2.6) are equivalent to assuming that uncorrelated zero-mean white noises with variances $\tilde{\sigma}_y$ and $\tilde{\sigma}_u$ are added to the input-output sequences.

Fig. 1 shows the measurement scheme, resulting from conditions (2.4), which can

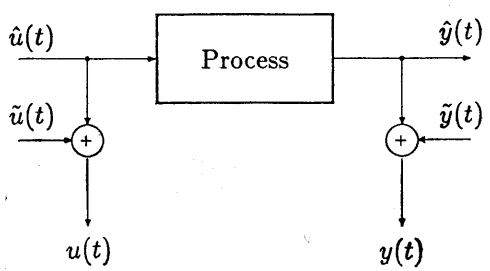

Fig. 1. Structure for identification.

describe many real situations; the available data are $u(\cdot)$ and $y(\cdot)$ only, and the goal of the identification consists in determining linear dynamic relations among the noiseless terms $\hat{u}(\cdot)$ and $\hat{y}(\cdot)$. With such assumptions solution of the identification problem consists in determining a value of k and a non-negative definite matrix $\tilde{\Sigma}_k = \text{diag}\,[\tilde{\sigma}_y I_{k+1}, \tilde{\sigma}_u I_{k+1}]$ such that

$$\tilde{\Sigma}_k = \Sigma_k - \text{diag}\,[\tilde{\sigma}_y I_{k+1}, \tilde{\sigma}_u I_{k+1}] \geqq 0\,. \tag{2.7}$$

Considering, on the noisy matrix Σ_k, the same partition (2.3) as $\hat{\Sigma}_k$, the following results holds [5].

Theorem 2.1. For a given value of k, the solution set of relation (2.7) describes, in the first quadrant of the $(\tilde{\sigma}_y, \tilde{\sigma}_u)$-plane, a continuous curve whose concavity faces the origin. For every $\tilde{\sigma}_u$ with $0 \leqq \tilde{\sigma}_u \leqq \tilde{\sigma}_u^k$, where

$$\tilde{\sigma}_u^k = \min \text{eig}\,\{\Sigma_k(uu) - \Sigma_k(uy)\,\Sigma_k^{-1}(yy)\,\Sigma_k^{\mathrm{T}}(uy)\} \tag{2.8}$$

the corresponding value $\tilde{\sigma}_y$ on the curve is given by

$$\tilde{\sigma}_y = \min \text{eig}\,\{\Gamma_k\} \tag{2.9a}$$

where

$$\Gamma_k = \Sigma_k(yy) - \Sigma_k^{\mathrm{T}}(uy)\,[\Sigma_k(uu) - \tilde{\sigma}_u I_{k+1}]^{-1}\,\Sigma_k(uy) \tag{2.9b}$$

and min eig $\{\cdot\}$ denotes the minimal eigenvalue of $\{\cdot\}$.

Proof. The proof is based on the following considerations. Since $\Sigma_k(yy)$ is non-singular the symmetric matrix $\Sigma_k - \text{diag}\,[0 I_{k+1}, \tilde{\sigma}_u^k I_{k+1}]$ is equivalent, as a quadratic form, to

$$\text{diag}\,[\Sigma_k(yy), \Sigma_k(uu) - \Sigma_k(uy)\,\Sigma_k^{-1}(yy)\,\Sigma_k^{\mathrm{T}}(uy) - \tilde{\sigma}_u^k I_{k+1}] \tag{2.10}$$

Since $\Sigma_k(yy)$ is positive definite then condition (2.7) is satisfied if and only if $\Sigma_k(uu) - \Sigma_k(uy)\,\Sigma_k^{-1}(yy)\,\Sigma_k^{\mathrm{T}}(uy) - \tilde{\sigma}_u^k I_{k+1} \geqq 0$, i.e. only when $\tilde{\sigma}_u^k$ satisfies relation (2.8). In a similar fashion for every $\tilde{\sigma}_u$ with $0 \leqq \tilde{\sigma}_u \leqq \tilde{\sigma}_u^k$ the corresponding value $\tilde{\sigma}_y$ satisfying relation (2.7) can be computed as in (2.9). □

Note that relation (2.7) is satisfied, for every value of k, by an infinity of solutions and that such solutions are not discriminable in the context of the proposed scheme. Consequently, the search for a model describing the available data, on the basis of an assumed system order, does not lead to a single solution, as happens in the algebraic case [3].

However, considering sets of solutions associated to increasing values of k, the following result can be obtained [5].

Theorem 2.2. For every $\tilde{\sigma}_u$ with $0 \leqq \tilde{\sigma}_u \leqq \tilde{\sigma}_u^{k+1}$ we have

$$\min \text{eig}\,\{\Gamma_{k+1}\} \leqq \min \text{eig}\,\{\Gamma_k\}\,. \tag{2.11}$$

Loosely speaking, the curves associated to increasing values of k approach, in the noise plane, the origin, and every curve includes all subsequent ones. Condition (2.11)

allows, however, possible common points between different curves or even their coincidence.

It is important to observe that, if the system generating the data belongs to the model set parametrized by equation (2.1) and the assumptions on the noise are satisfied, all the curves with $k \geqq n$ have necessarily at least one common point, i.e. the point $(\tilde{\sigma}_y^*, \tilde{\sigma}_u^*)$ corresponding to the true noise values

$$\tilde{\sigma}_y^* = \frac{1}{L+1} \sum_{i=0}^{L} \tilde{y}(i)^2 , \quad \tilde{\sigma}_u^* = \frac{1}{L+1} \sum_{i=0}^{L} \tilde{u}(i)^2 . \tag{2.12}$$

Note that for $L \to \infty$ the values $\tilde{\sigma}_y^*$ and $\tilde{\sigma}_u^*$ tend to the true variances of the white noises affecting the output and the input of the system.

It can be stated that the considered system is univocally identifiable if there exists an integer k^* such that all the curves (2.9) associated to values of k larger than k^* have one common point only. In this case $k^* = n$ and the parameter vector p can be computed as a basis of the null space of $\hat{\Sigma}_n$ with the $(n+1)$-th entry normalized to one.

Note that three different aspects influence the uniqueness of the solution: the structure of the considered model; the properties of the noise affecting the data; the input signal characteristics. For example, the reader can easily evaluate what happens with the system $\hat{y}(t+n) = \alpha \hat{u}(t)$, with $n > 0$, when the input $\hat{u}(\cdot)$ exhibits the same characteristics as the noises $\tilde{u}(\cdot)$ and $\tilde{y}(\cdot)$, that is, $\hat{\Sigma}_k(\hat{u}, \hat{u}) = \hat{\sigma}_u I_{k+1}$. In this case the curves associated to values of $k \geqq n$ are coincident segments of rectangular hyperbola and α cannot be univocally estimated. In any case, by selecting a suitable input sequence it is possible to identify the noise terms and, consequently, the order and parameters of the system.

The previous considerations allow us to conclude that, unlike the algebraic case, the Frisch scheme applied to dynamic system identification generally leads to a single solution. The additional information necessary to obtain this single model is carried by the correlations established among the samples by the dynamic nature of the process.

3. APPROXIMATE MODELLING OF REAL DYNAMIC PROCESSES

When real processes are considered, the hypotheses of linearity, time-invariance and finite-dimensionality are in general violated, the scheme previously assumed for the noise may be unrealistic and the available data set may be not complete. However, the search for a model satisfying the previous assumptions can be useful also in this case in order to obtain with little computational effort an approximate solution of limited complexity, which can be effective in several control engineering applications. The term noise loses in this context any stochastic meaning and can be simply considered as any deviation from the behavior of the assumed model.

If the proposed Frisch scheme is applied to real data, no common point among the curves associated to different model orders can be expected, since no exact model describing the data exists. The steps previously described will lead to a set of curves which do not cross each other; every curve encompasses all subsequent ones.

For a given order k, that is for a given model complexity, the infinite points of the corresponding curve are associated to an infinite number of solutions. A physical interpretation of the different points belonging to the same curve is suggested by the following property.

Remark 3.1. Two different points on the same curve characterize only two different decompositions of the matrix $\Sigma_k = \hat{\Sigma}'_k + \tilde{\Sigma}'_k = \hat{\Sigma}''_k + \tilde{\Sigma}''_k$, where $\hat{\Sigma}'_k$ and $\hat{\Sigma}''_k$ yield the models M' and M'' and $\tilde{\Sigma}'_k$, $\tilde{\Sigma}''_k$ are linked to the noise terms $(\tilde{\sigma}'_u, \tilde{\sigma}'_y)$ and $(\tilde{\sigma}''_u, \tilde{\sigma}''_y)$ respectively.

A property of the models associated to points belonging to the same curve is described by the following theorem.

Theorem 3.1. Models M' and M'' as defined in Remark 3.1 are not input-output equivalent.

Proof. Two irreducible input-output equivalent models are parametrized by the same vector $p = [\alpha_0, \ldots, \alpha_{k-1}, 1, \beta_0, \ldots, \beta_k]^T$, which must satisfy the relations $(\Sigma_k - \tilde{\Sigma}'_k)\, p = 0$ and $(\Sigma_k - \tilde{\Sigma}''_k)\, p = 0$, where $\tilde{\Sigma}'_k = \operatorname{diag}[\tilde{\sigma}'_y I_{k+1}, \tilde{\sigma}'_u I_{k+1}]$ and $\tilde{\Sigma}''_k = \operatorname{diag}[\tilde{\sigma}''_y I_{k+1}, \tilde{\sigma}''_u I_{k+1}]$. From the previous conditions it follows that $(\tilde{\Sigma}'_k - \tilde{\Sigma}''_k)\, p = 0$, and, since $\operatorname{rank}[\tilde{\Sigma}'_k - \tilde{\Sigma}''_k] = 2(k+1)$, the unique trivial solution $p = 0$ is obtained. □

It must be noted that for any model of order k, associated to the noise values $(\tilde{\sigma}'_u, \tilde{\sigma}'_y)$, a model of order $k+1$ associated to the noise values $(\tilde{\sigma}''_u, \tilde{\sigma}''_y)$, with $\tilde{\sigma}''_u < \tilde{\sigma}'_u$ and/or $\tilde{\sigma}''_y < \tilde{\sigma}'_y$, can always be found.

In other words, as the model order increases the amount of noise assumed in the data decreases. This consideration leads to say that models with greater complexity can fit the available data $u(\cdot)$ and $y(\cdot)$ with better accuracy, even if this does not correspond to the physical behavior of the process.

The results of real process identification obtained with the described procedure frequently show that the curves associated to values of k larger than a certain k^* are close to each other. This behavior shows that the noises affecting the data do not decrease significantly when models with order larger than k^* are considered, and can therefore lead to a criterion for the selection of a suitable model order.

As already said, the Frisch scheme applied to the identification of real dynamic systems does not lead to a single solution. A subset of the family of admissible models can be selected if further information about the process generating the data is available or some additional assumptions (*prejudices*) are introduced.

From this point of view, it must be noted that some classical identification approaches are particular cases of this scheme. In fact, it can be observed that the

points in the noise plane obtained by intersecting the curves with the straight line bisecting the first quadrant, correspond to the solutions given by the eigenvector method, when increasing model orders are tested. This scheme is equivalent to assuming that the input and the output are corrupted by the same amount of noise.

Moreover, the points of the curves on the y-axis are the solutions corresponding to output error models and in this case the noise is assumed as present only on the output sequence.

The set of admissible solutions may be further restricted if assumptions on the process are added. For example, if the following conditions are included

– the model order cannot be larger than a given value k,

– the noise terms $\tilde{u}(t)$ and $\tilde{y}(t)$ $(t = 0, \ldots, k + L)$ are bounded by relations of the type

$$-\tilde{u}_{\max} \leqq \tilde{u}(t) \leqq \tilde{u}_{\max}\,, \quad -\tilde{y}_{\max} \leqq \tilde{y}(t) \leqq \tilde{y}_{\max}\,, \tag{3.1}$$

then the identification problem reduces to determining a model with limited complexity, in the bounded subset of the noise plane

$$0 \leqq \tilde{\sigma}_u \leqq \tilde{u}^2_{\max}\,, \quad 0 \leqq \tilde{\sigma}_y \leqq \tilde{y}^2_{\max}\,. \tag{3.2}$$

When the hypotheses of the scheme are only slightly violated (e.g. weakly time-dependent or weakly non-linear processes), the behavior of the actual curves exhibits only a limited deviation from the ideal case.

If a single model is the goal of the identification procedure, it can be reasonable to look for the point on the curve corresponding to the selected order that minimizes the total amount of noise affecting the input-output data, defined as $\sqrt{\tilde{\sigma}_u^2 + \tilde{\sigma}_y^2}$.

This selection rule, however, is not consistent since it does not lead necessarily to the exact solution when the Frisch scheme assumptions are fulfilled by the process.

Recalling the main feature of the point representing the exact solution in the ideal case (i.e. this point belongs to all the curves associated to orders larger than the system one), the search for the point on the curve of the selected order k nearest to the subsequent curve can be considered a consistent criterion to obtain a single solution when only small violations of the Frisch scheme assumptions are present.

A general strategy can thus consist in a preliminary evaluation of the compatibility of the available sequences with the hypotheses of the proposed scheme, on the basis of the minimal distance between the curve associated to the selected model order and the subsequent one. If compatibility exists the latter criterion is more suitable since the modelling problem can be seen as a search for the "true" system generating the data. Note that the model obtained can also be far from the process under study because of lack of completeness in the data; when the data are complete, however, this model gives a consistent description of both the data and the process.

When the compatibility between the data and the scheme is not fulfilled, then the first criterion (noise minimization) can be preferred and the modelling problem can be considered as the approximation of the available data by means of a linear model in the context of the Frisch scheme.

By choosing the Euclidean norm as metric in the noise space $(\tilde{\sigma}_u, \tilde{\sigma}_y)$, the points with minimal distance between two consecutive curves associated to orders k and $k + 1$ can be obtained as follows.

Algorithm 3.1.

- Every point $P''(\tilde{\sigma}_u'', \tilde{\sigma}_y'')$ on the curve associated to the order $k + 1$ can be computed by using the algorithm described in Theorem 2.1;
- the search for the point on the curve of order k at minimal distance from P'' is restricted to the points $P'(\tilde{\sigma}_u'' + \Delta\tilde{\sigma}_u, \tilde{\sigma}_y'' + \Delta\tilde{\sigma}_y)$, with positive values $(\Delta\tilde{\sigma}_u, \Delta\tilde{\sigma}_y)$ given by relation

$$\begin{aligned}\hat{\Sigma}_k &= \Sigma_k - (\operatorname{diag}[\tilde{\sigma}_y'' I_{k+1}, \tilde{\sigma}_u'' I_{k+1}] + \operatorname{diag}[\Delta\tilde{\sigma}_y I_{k+1}, \Delta\tilde{\sigma}_u I_{k+1}]) \\ &= \Sigma_k' - \operatorname{diag}[\Delta\tilde{\sigma}_y I_{k+1}, \Delta\tilde{\sigma}_u I_{k+1}] \geqq 0 . \end{aligned} \tag{3.3}$$

  This is a Frisch scheme problem of type (2.7) for matrix Σ_k'. The last equality in (3.3) implies a translation of the origin of the noise space to the point $P''(\tilde{\sigma}_u'', \tilde{\sigma}_y'')$. The minimal distance can be denoted with $D_{P''P'^*} = \sqrt{(\Delta\tilde{\sigma}_u^*)^2 + (\Delta\tilde{\sigma}_y^*)^2}$, where P'^* is the point $(\tilde{\sigma}_u'' + \Delta\tilde{\sigma}_u^*, \tilde{\sigma}_y'' + \Delta\tilde{\sigma}_y^*)$;
- The minimum value of $D_{P''P'^*}$ when the point P'' describes the whole curve of order $k + 1$ is the minimal distance $D_{P''^*P'^*}$ between the curves of order k and $k + 1$.

This distance can be considered a measure of the perturbation on the filtering matrix $\tilde{\Sigma}_k$ associated to the point P'^* necessary to obtain a noise model compatible with a solution of order $k + 1$. For the chosen set of models with order k, such distance describes in mathematical terms the validity of the assumptions of the Frisch scheme for the available data.

Note that previous considerations refer to a noise plane whose axes report absolute measures of the noise terms $\tilde{\sigma}_u$ and $\tilde{\sigma}_y$. However, it must be noted that when all the samples of the input (output) sequence are multiplied by a constant $M_u(M_y)$ (e.g. choice of different units) the values of $\tilde{\sigma}_u(\tilde{\sigma}_y)$ are multiplied by $M_u^2(M_y^2)$. To achieve identification results independent of input (output) scaling, the following per cent measures can be introduced

$$\tilde{\sigma}_{u\%} = 100 \frac{\dfrac{1}{L+1} \sum_{i=0}^{L} \tilde{u}(i)^2}{\dfrac{1}{N} \sum_{i=0}^{N-1} u(i)^2} \tag{3.4a}$$

$$\tilde{\sigma}_{y\%} = 100 \frac{\dfrac{1}{L+1} \sum_{i=0}^{L} \tilde{y}(i)^2}{\dfrac{1}{N} \sum_{i=0}^{N-1} y(i)^2} . \tag{3.4b}$$

They represent the per cent sampled variances of the estimated additive noise affecting the input-output sequences, normalized with respect to the available data. Note that the numerator sum in (3.4) refers to $L + 1$ samples since this is the number of terms used to construct the diagonal elements of Σ_k, while the denominator sum refers to the whole number of available samples.

Algorithm 3.1 thus leads to an absolute minimal distance point in the plane $(\tilde{\sigma}_u\%, \tilde{\sigma}_y\%)$ if the input and output data have been normalized by dividing each sample by $\sqrt{(N^{-1}\sum_{i=0}^{N-1} u(i)^2)}$ and $\sqrt{(N^{-1}\sum_{i=0}^{N-1} y(i)^2)}$ respectively.

Example 3.1. Let us consider a single-input single-output simulated system with a state space model of order 3 and with a non-linear term given by the product of a state component by the input of the system. The system has been excited by a pseudo random input sequence and the input-output data have been corrupted by additive white noises with per cent sampled variances of type (3.3) $\tilde{\sigma}_u\% = 20$ and $\tilde{\sigma}_y\% = 10$ respectively. The previous point has been denoted with a star in Fig. 2 which shows the admissible solutions in the noise space, when the process is identified by means of a linear model. The outmost curve refers to models with order one while the inmost one refers to models with order 5. It can be noted that curves corresponding to an order larger than 3 are very close to each other and, consequently, the choice of a model with order 3 can be considered as reasonable.

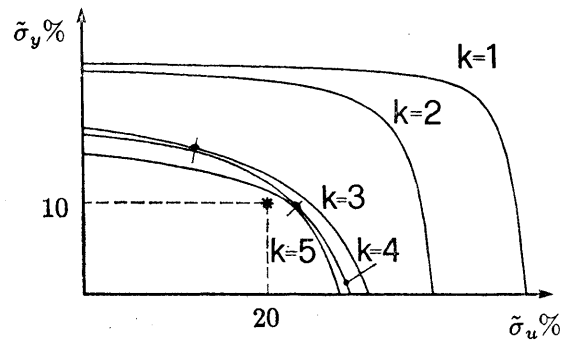

Fig. 2. Curves of admissible solutions in the noise space.

Figure 2 shows the segment of minimal distance between the curves of order 3 and 4; it gives a measure of the assumption validity of the Frisch scheme when a model with order 3 is chosen. Note that the selection of a model with order 4 leads to similar considerations for the assumption evaluation, but to different indications on the input-output noises. In both cases the results obtained differ, as expected, from the sampled variances of the actual noises affecting the non-linear system.

The identification approach previously described has been applied to real data regarding production/storage processes in natural gas reservoirs. The results have been described in [6] and are conceptually similar to those proposed here. In those

applications the procedure has shown that the available data could be explained by linear models of low order with a small amount of additive noises. Since all the models of a given order exhibit limited parameter scattering, the search for a single solution plays, in these cases, a secondary role.

4. CONCLUSIONS

This paper briefly describes the results obtained in the extension of the algebraic Frisch scheme to the identification of dynamic systems. This approach leads to the true solution when the process under study and the noises affecting the data satisfy the assumptions of the scheme, otherwise a whole set of admissible solutions is obtained.

When a real process is considered, a subset of such a family can be selected if further information about the process is available or additional assumptions, generally linked to the planned use of the identified models, are introduced. Once the identification has been performed it is important to analyze the validity of the stated assumptions.

This paper has introduced the concept of "distance" to evaluate how much the identified models violate the scheme hypotheses and a consistent criterion to select a single model among the infinite admissible ones. Finally, some results regarding a simulated system have been reported to explain the characteristics of the approach proposed.

(Received November 26, 1990.)

REFERENCES

[1] J. C. Willems: From time series to linear system, Part III. Automatica *23* (1987), 87—115.

[2] R. Frisch: Statistical confluence analysis by means of complete regression systems. Pub. No. 5 (1934), University of Oslo, Economics Institute.

[3] R. E. Kalman: System identification from noisy data. In: Dynamical Systems II (A.R. Bednarek and L. Cesari, ed.), Academic Press, New York 1982, pp. 135—164.

[4] M. Deistler: Linear errors-in-variables models. In: Time Series and Linear Systems (S. Bittanti, ed.; Lecture Notes in Control and Information Sciences 86), Springer-Verlag, Berlin—Heidelberg—New York—Tokyo 1984, pp. 37—68.

[5] S. Beghelli, R. P. Guidorzi and U. Soverini: The Frisch scheme in dynamic system identification. Automatica *26* (1990), 171—176.

[6] S. Beghelli, R. P. Guidorzi and U. Soverini: Data compatible models in gas storage identification. International Conference of System Science X, Wroclaw, Poland, September 1989.

Professor Sergio Beghelli, Professor Roberto P. Guidorzi, Dr. Ing. Umberto Soverini, Dipartimento di Elettronica, Informatica e Sistemistica, Università di Bologna, Viale del Risorgimento 2, I 40136 Bologna. Italy.