System Identification and Control: a fruitful cooperation over half a century, and more

31 CCC 2012

July 2012

Focus:
System Identification evolution in the realm of Control Science

Paradigm problem: parameter estimation

Outline

1) Foreword
Milano identification & control connections
- Beijing
- Laxenburg
- Moscow

2) The parameter estimation problem

3) Available estimation tools (50 years of research)
- PE
- KF

4) Proposal of a new method
- Two stage TS

5) Conclusion

1) Foreword:
Identification and control connection between Milano and …
- Beeijing - Laxenburg - Moscow

Beijing ⇨ Milano

Beijing → Milano

Laxenburg 1977

(after Düsseldorf 1957 – 1975
and Helsinki, 1975-1977)

Laxenburg ⇨ Milano

IFAC headquarters in Laxenburg

Brera building in Milano

Laxenburg ⇨ Milano

IFAC headquarters in Laxenburg

Brera building in Milano

control library

Control library in Laxenburg

Library in Brera building in Milano

Libertà Eguaglianza

IN NOME DELLA REPUBBLICA CISALPINA

ESTRATTO

DE' REGISTRI DEL DIRETTORIO ESECUTIVO

Seduta del giorno 19. Brumale anno VI. Repubblicano

Il Generale in Capo dell'Armata d'Italia in nome della Repubblica Francese ha fatto deporre presso il Direttorio Esecutivo la seguente

LEGGE

Considerando, che a norma dell'Art. 297. della Costituzione vi deve essere per tutta la Repubblica un Istituto Nazionale incaricato di raccogliere le scoperte, e perfezionare le Arti, e le Scienze;
Considerando ancora, che ampli ed opportuni stabilimenti utili a questo oggetto distinguono specialmente la Comune di Bologna

SI DECRETA

L'Istituto Nazionale della Repubblica Cisalpina è fissato in Bologna.

FONTANA = BRAGALDI = ALDINI = SEVEROLI = RE = GUICCIARDI
MASCHERONI = VERTEMATE FRANCHI = LOSCHI = GELMETTI

Approuvé = BONAPARTE

Il Direttorio Esecutivo ordina che la premessa Legge sia munita del Sigillo della Repubblica, pubblicata, ed eseguita.

ALESSANDRI PRESIDENTE

LUOGO DEL SIGILLO

Pel Direttorio Esecutivo
Il Segretario Generale
SOMMARIVA

Per copia conforme

IL MINISTRO DELLA GIUSTIZIA, E DELLA POLIZIA GENERALE

LUOSI

Bellerio *Segretario.*

Per li Fratelli Pirola Impressori della Repubblica. Si vende un soldo.

Istituto Lombardo –
Milano Academy of Science and Literature

The IFAC - Khane donation

March 24 2011
– 18 boxes full of IFAC Proceedings
at Lombardo
- donation of Steve Kahne (pres. 93-96)

The Khane donation to Lombardo

The Khane donation to Lombardo

Visiting Lombardo (IFAC 2011, 1st Sept.)

Visiting Lombardo (IFAC 2011, 1st Sept.)

Visiting Lombardo (IFAC 2011, 1st Sept.)

Call for books…

missing items

1st World Congress
2nd World Congress
…

Call for books…

Would you like to complete the collection?

Moscow ⇨ Milano

1960
IFAC World Congress (Moscow)

2011
IFAC World Congress (Milan)

3629 papers
7140 authors
73 nations

Website
ifac2011.org
active
(page-rank 6)

2011 – 18th IFAC World Congress – Milan: *Historical Session*

S. Kahne, R.E. Kalman, M. Thoma, T. Vamos, J. Westcott
(A. Kurzhanski chair)

2011 – 18th IFAC World Congress – Milan: *Historical Session*

“In any engineering application the success of control is determined by the accuracy with which the model represents data,
and this leads to the realm of Identification”

Rudolph Kalman

1960 – 1st IFAC World Congress - Moscow

On the General Theory of Control Systems

R. E. KALMAN

Introduction

In no small measure, the great technological progress in automatic control and communication systems during the past two decades has depended on advances and refinements in the mathematical study of such systems. Conversely, the growth of technology brought forth many new problems (such as those related to using digital computers in control, etc.) to challenge the ingenuity and competence of research workers concerned with theoretical questions.

Despite the appearance and effective resolution of many new problems, our understanding of fundamental aspects of control has remained superficial. The only basic advance so far appears to be the *theory of information* created by Shannon[1]. The chief significance of his work in our present interpretation is the discovery of general 'laws' underlying the process of information transmission, which are quite independent of the particular models being considered or even the methods used for the description and analysis of these models. These results could be compared with the 'laws' of physics, with the crucial difference that the 'laws' governing man-made objects cannot be discovered by straightforward experimentation but only by a purely abstract analysis guided by intuition gained in observing present-day examples of technology and economic organization. We may thus classify Shannon's result as belonging to the *pure theory* of communication and control, while everything else can be labelled as the *applied theory*; this terminology reflects the well-known distinctions between pure and applied physics or mathematics. For reasons pointed out above, in its methodology the pure theory of communication and control closely resembles mathematics, rather than physics; however, it is not a branch of mathematics because at present we cannot (yet?) disregard questions of physical realizability in the study of mathematical models.

This paper initiates study of the pure theory of control, imitating the spirit of Shannon's investigations but otherwise using entirely different techniques. Our ultimate objective is to answer questions of the following type: What kind and how much information is needed to achieve a desired type of control? What intrinsic properties characterize a given unalterable plant as far as control is concerned?

At present only superficial answers are available to these questions, and even then only in special cases.

Initial results presented in this Note are far from the degree of generality of Shannon's work. By contrast, however, only *constructive* methods are employed here, giving some hope of being able to avoid the well-known difficulty of Shannon's theory: methods of proof which are impractical for actually constructing practical solutions. In fact, this paper arose from the need for a better understanding of some recently discovered computation methods of control-system synthesis[2-5]. Another by-product of the paper is a new computation method for the solution of the classical Wiener filtering problem[7].

The organization of the paper is as follows:

In Section 3 we introduce the models for which a fairly complete theory is available: dynamic systems with a finite dimensional state space and linear transition functions (i.e. systems obeying linear differential or difference equations). The class of random processes considered consists of such dynamic systems excited by an uncorrelated gaussian random process. Other assumptions, such as stationarity, discretization, single input/single output, etc., are made only to facilitate the presentation and will be absent in detailed future accounts of the theory.

In Section 4 we define the concept of *controllability* and show that this is the 'natural' generalization of the so-called 'dead-beat' control scheme discovered by Oldenbourg and Sartorius[21] and later rederived independently by Tsypkin[22] and the author[17].

We then show in Section 5 that the general problem of optimal regulation is solvable if and only if the plant is completely controllable.

In Section 6 we introduce the concept of *observability* and solve the problem of reconstructing unmeasurable state variables from the measurable ones in the minimum possible length of time.

We formalize the similarities between controllability and observability in Section 7 by means of the *Principle of Duality* and show that the Wiener filtering problem is the natural dual of the problem of optimal regulation.

Section 8 is a brief discussion of possible generalizations and currently unsolved problems of the pure theory of control.

Notation and Terminology

The reader is assumed to be familiar with elements of linear algebra, as discussed, for instance, by Halmos[8].

Consider an n-dimensional real vector space X. A *basis* in X is a set of vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ in X such that any vector $\mathbf{x}$ in X can be written uniquely as

$$\mathbf{x} = x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n \qquad (1)$$

the x_i being real numbers, the *components* or *coordinates* of $\mathbf{x}$. Vectors will be denoted throughout by small bold-face letters.

The set X^* of all real-valued linear functions $\mathbf{x}^*$ (= *covectors*) on X, with the 'natural' definition of addition and scalar multiplication, is an n-dimensional vector space. The value of a covector $\mathbf{y}^*$ at any vector $\mathbf{x}$ is denoted by $[\mathbf{y}^*, \mathbf{x}]$. We call this the *inner product* of $\mathbf{y}^*$ by $\mathbf{x}$. The vector space X^* has a natural basis $\mathbf{a}^*_1, \dots, \mathbf{a}^*_n$ associated with a given basis in X; it is defined by the requirement that

$$[\mathbf{a}^*_i, \mathbf{a}_j] = \delta_{ij} \qquad (2)$$

Using the 'orthogonality relation' 2, we may write 1 in the form

$$\mathbf{x} = \sum_{i=1}^{n} [\mathbf{a}^*_i, \mathbf{x}]\mathbf{a}_i \qquad (3)$$

which will be used frequently.

For purposes of numerical computation, a vector may be considered a matrix with one column and a covector a matrix

- Paradigm Shift
- Control science & technology becomes more and more a model based discipline

1960 – 1st IFAC World Congress - Moscow

The Parameter Estimation Problem

J. H. WESTCOTT

Introduction

The establishment of causal relationships between variables forms the fundamental basis of the scientific method. It could be described as the task of devising mathematical models to predict, to known accuracy, the results of given actions or environments. In this paper systems of interacting quantities are considered where it is required to establish a mathematical expression for the interaction between variables. From observations of the variables over a period of time it is required to specify the values of parameters that best account for the observed behaviour or that will best serve to predict the system behaviour at future times.

The theory of control systems is essentially the theory of interacting quantities, but in addition they have the distinctive feature of a closed sequence of control resulting from the use of feedback associated with the possibilities of self-excitation and oscillation. The closed loop nature of such systems somewhat complicates the problem of distinguishing 'cause' from 'effect' in the records of the system variables which is of vital importance in a realistic identification of causal relationships. In many engineering control systems this problem does not arise since the system can be taken apart and the action of one quantity on another accurately determined in isolation from the rest of the system, but this is by no means always the case and in many interesting cases, such as in an automatically controlled chemical plant or an aircraft in flight, it may be technically difficult or economically impossible to measure experimentally the separate dependencies that constitute the system.

When using deterministic inputs as engineers are accustomed to doing, the input is specified with absolute precision and so it is permissible to construct an impressive edifice of precise logical operations on the original signal; but in the realm of probability theory facts are not so precisely defined and one must be satisfied with a more humble statement of conclusions consonant with the uncertainties concerning the input. It is insufficient to draw conclusions under these circumstances without also accompanying them with a statement as to their accuracy in the light of the statistical nature of the inputs.

Parameter Estimation

It is necessary to distinguish clearly between two distinct approaches to the parameter estimation problem. In the first approach, which is the simpler of the two, a causal relationship is assumed to exist between input and output quantities, in which the relationship is obscured by the presence of a random disturbance. The problem is then treated as a data-fitting one in which parameters are chosen, which for the particular input and output given attributes minimum effect to the disturbance.

The procedure is one that will be familiar as least-squares fitting of data[1] and involves the reduction of a redundant set of equations to a set of simultaneous equations in which sums of products of pairs of variables replace particular values of the variables. The method can be extended to deal with data expressed as continuous time functions in which case the sums of products become integrals of products taken over the duration of the records. These integrals have a superficial resemblance to the following expression for the correlation function:

$$\phi_{ff}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} f(t) \,.\, f(t+\tau) \,.\, \mathrm{d}t \qquad (1)$$

The important difference is that they do not extend to the limit as time approaches infinity, and there is no reference to the properties of the hypothetical population of records from which the parameter records may be regarded as selected.

The second method of parameter estimation involves taking a broader statistical view of the same problem, that is to say it is assumed that stochastic terms have statistical regularity that is governed by a fixed set of statistical parameters and probability distributions. This means that the variables in the system will be statistically fluctuating time series, and any measured record may be considered to be one of many possible ones. In general, a further record taken from the system will give different sets of parameter values. Thus what is required is to calculate from the given record the best approximation to the underlying statistical parameters by means of which the variance of estimates of the parameter values may be deduced.

Much work has been done by engineers which is referred to in the next section and can be utilized to discover causal relationships between variables when one of the variables is assumed to be associated with a random disturbance, but this work has been based on the assumption of an infinitely long record. While this is valuable there is still a problem of allowing for the uncertainty that exists when conclusions must be drawn from a record of finite duration, as the outcome of a practical experiment. In general, the longer the duration of the experiment the more accurate will be the statements that can be made concerning the relationships, but these statements will also be influenced by the degree to which the system concerned has discernible stationary statistical parameters.

The existence of statistical regularity in the records makes it possible to make statements concerning the probable behaviour of the system in a future experiment. In some cases it is possible to attach confidence limits to such predictions. This is the crux of the problem of drawing conclusions from short records. The brevity of a record in this context is not decided by its actual duration. It is decided by the minimum period for which the underlying statistical parameters may be considered as stationary. Thus if a record is defined as short within our present meaning we are asserting our belief that this minimum period is comparable with its duration and that we consider it appropriate to make statistical inference from the record. If the purpose for which the analysis is being made is the design of a model for the system from which the records were derived, then the model is required to give the best possible representation of the system under conditions which are statistically

- John Westcott: The Parameter Estimation Problem

1960 – 1st IFAC World Congress - Moscow

The Parameter Estimation Problem

J. H. WESTCOTT

Introduction

the variables. The method can be extended to deal with data

On the General Theory of Control Systems

R. E. KALMAN

Introduction

In no small measure, the great technological progress in automatic control and communication systems during the past two decades has depended on advances and refinements in the mathematical study of such systems. Conversely, the growth of technology brought forth many new problems (such as those related to using digital computers in control, etc.) to challenge the ingenuity and competence of research workers concerned with theoretical questions.

Despite the appearance and effective resolution of many new problems, our understanding of fundamental aspects of control has remained superficial. The only basic advance so far appears to be the *theory of information* created by Shannon[1]. The chief significance of his work in our present interpretation is the discovery of general 'laws' underlying the process of information transmission, which are quite independent of the particular models being considered or even the methods used for the description and analysis of these models. These results could be compared with the 'laws' of physics, with the crucial difference that the 'laws' governing man-made objects cannot be discovered by straightforward experimentation but only by a purely abstract analysis guided by intuition gained in observing present-day examples of technology and economic organization. We may thus classify Shannon's result as belonging to the *pure theory* of communication and control, while everything else can be labelled as the *applied theory*; this terminology reflects the well-known distinctions between pure and applied physics or mathematics. For reasons pointed out above, in its methodology the pure theory of communication and control closely resembles mathematics, rather than physics; however, it is not a branch of mathematics because at present we cannot (yet?) disregard questions of physical realizability in the study of mathematical models.

This paper initiates study of the pure theory of control, imitating the spirit of Shannon's investigations but otherwise using entirely different techniques. Our ultimate objective is to answer questions of the following type: What kind and how much information is needed to achieve a desired type of control? What intrinsic properties characterize a given unalterable plant as far as control is concerned?

At present only superficial answers are available to these questions, and even then only in special cases.

Initial results presented in this Note are far from the degree of generality of Shannon's work. By contrast, however, only *constructive* methods are employed here, giving some hope of being able to avoid the well-known difficulty of Shannon's theory: methods of proof which are impractical for actually constructing practical solutions. In fact, this paper arose from the need for a better understanding of some recently discovered computation methods of control-system synthesis[2-5]. Another by-product of the paper is a new computation method for the solution of the classical Wiener filtering problem[7].

The organization of the paper is as follows:

In Section 3 we introduce the models for which a fairly complete theory is available: dynamic systems with a finite dimensional state space and linear transition functions (i.e. systems obeying linear differential or difference equations). The class of random processes considered consists of such dynamic systems excited by an uncorrelated gaussian random process. Other assumptions, such as stationarity, discretization, single input/single output, etc., are made only to facilitate the presentation and will be absent in detailed future accounts of the theory.

In Section 4 we define the concept of *controllability* and show that this is the 'natural' generalization of the so-called 'dead-beat' control scheme discovered by Oldenbourg and Sartorius[21] and later rederived independently by Tsypkin[22] and the author[17].

We then show in Section 5 that the general problem of optimal regulation is solvable if and only if the plant is completely controllable.

In Section 6 we introduce the concept of *observability* and solve the problem of reconstructing unmeasurable state variables from the measurable ones in the minimum possible length of time.

We formalize the similarities between controllability and observability in Section 7 by means of the *Principle of Duality* and show that the Wiener filtering problem is the natural dual of the problem of optimal regulation.

Section 8 is a brief discussion of possible generalizations and currently unsolved problems of the pure theory of control.

Notation and Terminology

The reader is assumed to be familiar with elements of linear algebra, as discussed, for instance, by Halmos[8].

Consider an n-dimensional real vector space X. A *basis* in X is a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in X such that any vector $\mathbf{x}$ in X can be written uniquely as

$$\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n \quad (1)$$

the x_i being real numbers, the *components* or *coordinates* of $\mathbf{x}$. Vectors will be denoted throughout by small bold-face letters.

The set X^* of all real-valued linear functions $\mathbf{x}^*$ (= *covectors*) on X, with the 'natural' definition of addition and scalar multiplication, is an n-dimensional vector space. The value of a covector $\mathbf{y}^*$ at any vector $\mathbf{x}$ is denoted by $[\mathbf{y}^*, \mathbf{x}]$. We call this the *inner product* of $\mathbf{y}^*$ by $\mathbf{x}$. The vector space X^* has a natural basis $\mathbf{a}^*_1, \ldots, \mathbf{a}^*_n$ associated with a given basis in X; it is defined by the requirement that

$$[\mathbf{a}^*_i, \mathbf{a}_j] = \delta_{ij} \quad (2)$$

Using the 'orthogonality relation' 2, we may write 1 in the form

$$\mathbf{x} = \sum_{i=1}^{n} [\mathbf{a}^*_i, \mathbf{x}]\mathbf{a}_i \quad (3)$$

which will be used frequently.

For purposes of numerical computation, a vector may be considered a matrix with one column and a covector a matrix

38

481

- Control Systems
- Parameter estimation
- Identification and control cooperating since long time

1960 – 1st IFAC World Congress - Moscow

The Parameter Estimation Problem

J. H. WESTCOTT

Introduction

the variables. The method can be extended to deal with data

On the General Theory of Control Systems

R. E. KALMAN

Introduction

In no small measure, the great technological progress in automatic control and communication systems during the past two decades has depended on advances and refinements in the mathematical study of such systems. Conversely, the growth of technology brought forth many new problems (such as those related to using digital computers in control, etc.) to challenge the ingenuity and competence of research workers concerned with theoretical questions.

Despite the appearance and effective resolution of many new problems, our understanding of fundamental aspects of control has remained superficial. The only basic advance so far appears to be the *theory of information* created by Shannon[1]. The chief significance of his work in our present interpretation is the discovery of general 'laws' underlying the process of information transmission, which are quite independent of the particular models being considered or even the methods used for the description and analysis of these models. These results could be compared with the 'laws' of physics, with the crucial difference that the 'laws' governing man-made objects cannot be discovered by straightforward experimentation but only by a purely abstract analysis guided by intuition gained in observing present-day examples of technology and economic organization. We may thus classify Shannon's result as belonging to the *pure theory* of communication and control, while everything else can be labelled as the *applied theory*; this terminology reflects the well-known distinctions between pure and applied physics or mathematics. For reasons pointed out above, in its methodology the pure theory of communication and control closely resembles mathematics, rather than physics; however, it is not a branch of mathematics because at present we cannot (yet?) disregard questions of physical realizability in the study of mathematical models.

This paper initiates study of the pure theory of control, imitating the spirit of Shannon's investigations but otherwise using entirely different techniques. Our ultimate objective is to answer questions of the following type: What kind and how much information is needed to achieve a desired type of control? What intrinsic properties characterize a given unalterable plant as far as control is concerned?

At present only superficial answers are available to these questions, and even then only in special cases.

Initial results presented in this Note are far from the degree of generality of Shannon's work. By contrast, however, only *constructive* methods are employed here, giving some hope of being able to avoid the well-known difficulty of Shannon's theory: methods of proof which are impractical for actually constructing practical solutions. In fact, this paper arose from the need for a better understanding of some recently discovered computation methods of control-system synthesis[2-5]. Another by-product of the paper is a new computation method for the solution of the classical Wiener filtering problem[7].

The organization of the paper is as follows:

In Section 3 we introduce the models for which a fairly complete theory is available: dynamic systems with a finite dimensional state space and linear transition functions (i.e. systems obeying linear differential or difference equations). The class of random processes considered consists of such dynamic systems excited by an uncorrelated gaussian random process. Other assumptions, such as stationarity, discretization, single input/single output, etc., are made only to facilitate the presentation and will be absent in detailed future accounts of the theory.

In Section 4 we define the concept of *controllability* and show that this is the 'natural' generalization of the so-called 'dead-beat' control scheme discovered by Oldenbourg and Sartorius[21] and later rederived independently by Tsypkin[22] and the author[17].

We then show in Section 5 that the general problem of optimal regulation is solvable if and only if the plant is completely controllable.

In Section 6 we introduce the concept of *observability* and solve the problem of reconstructing unmeasurable state variables from the measurable ones in the minimum possible length of time.

We formalize the similarities between controllability and observability in Section 7 by means of the *Principle of Duality* and show that the Wiener filtering problem is the natural dual of the problem of optimal regulation.

Section 8 is a brief discussion of possible generalizations and currently unsolved problems of the pure theory of control.

Notation and Terminology

The reader is assumed to be familiar with elements of linear algebra, as discussed, for instance, by Halmos[8].

Consider an n-dimensional real vector space X. A *basis* in X is a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in X such that any vector $\mathbf{x}$ in X can be written uniquely as

$$\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n \quad (1)$$

the x_i being real numbers, the *components* or *coordinates* of $\mathbf{x}$. Vectors will be denoted throughout by small bold-face letters.

The set X^* of all real-valued linear functions $\mathbf{x}^*$ (= *covectors*) on X, with the 'natural' definition of addition and scalar multiplication, is an n-dimensional vector space. The value of a covector $\mathbf{y}^*$ at any vector $\mathbf{x}$ is denoted by $[\mathbf{y}^*, \mathbf{x}]$. We call this the *inner product* of $\mathbf{y}^*$ by $\mathbf{x}$. The vector space X^* has a natural basis $\mathbf{a}^*_1, \ldots, \mathbf{a}^*_n$ associated with a given basis in X; it is defined by the requirement that

$$[\mathbf{a}^*_i, \mathbf{a}_j] = \delta_{ij} \quad (2)$$

Using the 'orthogonality relation' 2, we may write 1 in the form

$$\mathbf{x} = \sum_{i=1}^{n} [\mathbf{a}^*_i, \mathbf{x}]\mathbf{a}_i \quad (3)$$

which will be used frequently.

For purposes of numerical computation, a vector may be considered a matrix with one column and a covector a matrix

38

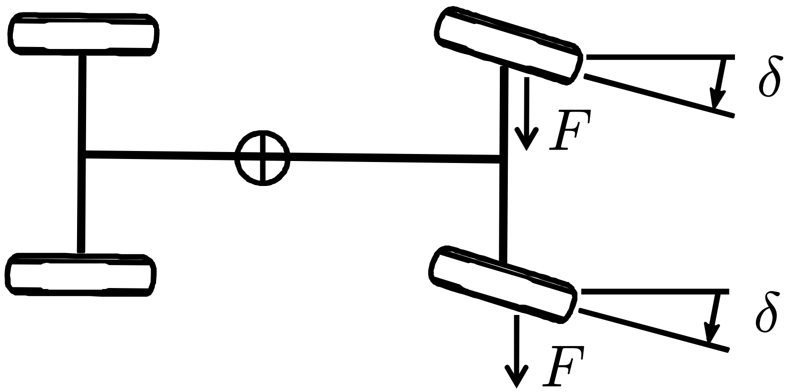
481

- Control Systems
- Parameter estimation
- Probing the progresses of estimation

2) The parameter estimation problem

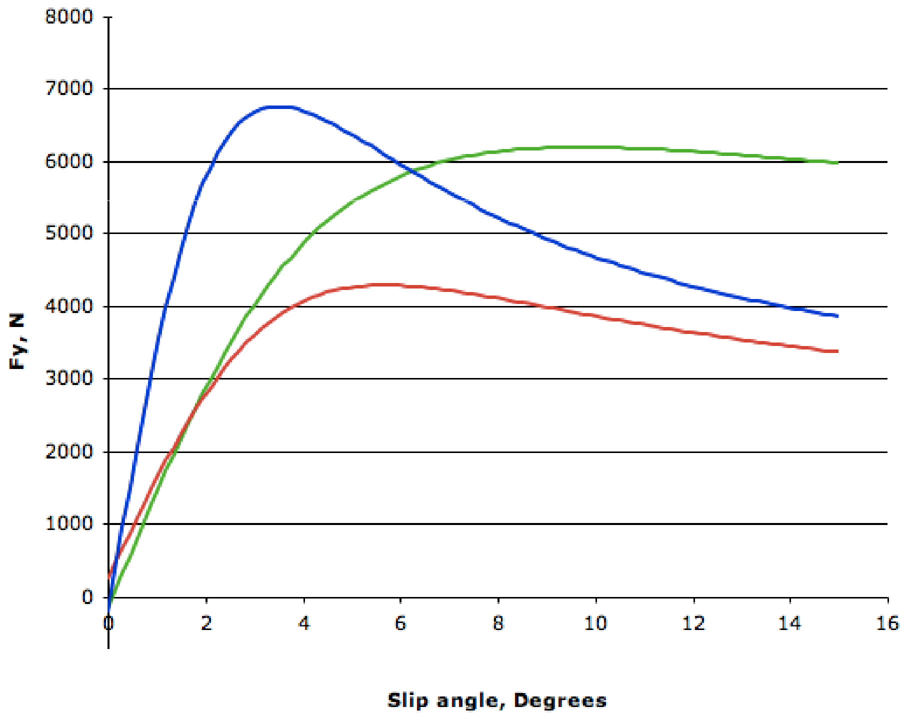
Parameter estimation in physical models

- Example: Tire



δ = steering angle

F = lateral force at the tires

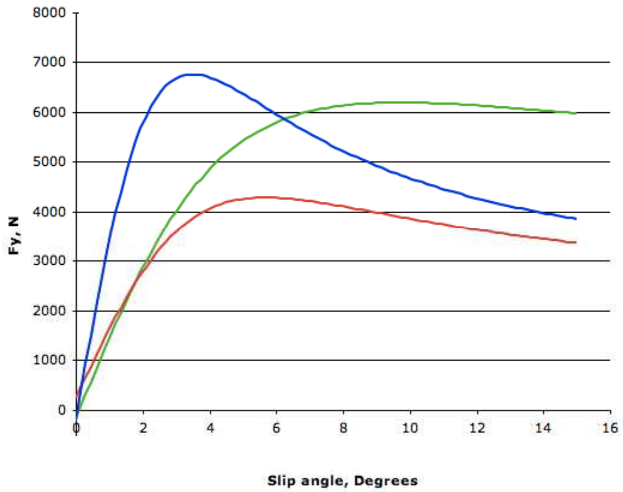


Parameter estimation in physical models

Pacejka's magic formula

$$F = D\sin\{C\arctan[B\bar{\alpha} - E(B\bar{\alpha} - \arctan(B\bar{\alpha}))]\} + S_V$$

$$\begin{aligned}
S_H &= p_{H1} + p_{H2} f_z + p_{H3}\gamma \\
S_V &= Q[(p_{V1} + p_{V2} f_z) + (p_{V3} + p_{V4} f_z)\gamma] \\
\bar{\alpha} &= \alpha(\delta) + S_H \\
C &= p_{C1} \\
\mu &= (p_{D1} + p_{D2} f_z)(1 - p_{D3}\gamma^2) \\
D &= \mu Q \\
E &= (p_{E1} + p_{E2} f_z)[1 - (p_{E3} + p_{E4}\gamma)\mathrm{sign}(\bar{\alpha})] \\
K &= p_{K1} F_z \sin[2\arctan(Q/(p_{K2} \cdot F_z)))(1 - p_{K3}|\gamma|)] \\
B &= K/(C \cdot D).
\end{aligned}$$

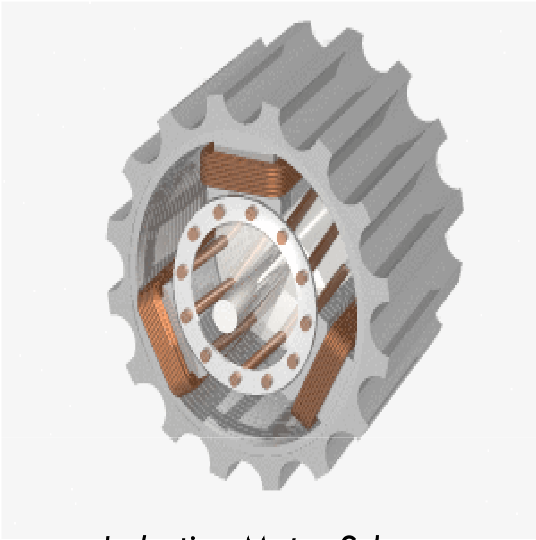


Pacejka's parameters to be estimated

Parameter estimation in physical models

- Example: Induction motor
- Critical Parameter: *rotor resistance* R_r

$$
\begin{aligned}
\frac{d\omega}{dt} &= \frac{M}{JL_r}(\psi_a i_b - \psi_b i_a) - \frac{T_l}{J} \\
\frac{d\psi_a}{dt} &= -\frac{R_r}{L_r}\psi_a - \omega\psi_b + \frac{R_r}{L_r}M i_a \\
\frac{d\psi_b}{dt} &= -\frac{R_r}{L_r}\psi_b + \omega\psi_a + \frac{R_r}{L_r}M i_b \\
\frac{di_a}{dt} &= -\left(\frac{R_s}{\sigma} + \frac{M}{\sigma L_r}\frac{R_r}{L_r}M\right)i_a + \frac{1}{\sigma}u_a + \frac{M}{\sigma L_r}\frac{R_r}{L_r}\psi_a + \frac{M}{\sigma L_r}\omega\psi_b \\
\frac{di_b}{dt} &= -\left(\frac{R_s}{\sigma} + \frac{M}{\sigma L_r}\frac{R_r}{L_r}M\right)i_a + \frac{1}{\sigma}u_b + \frac{M}{\sigma L_r}\frac{R_r}{L_r}\psi_b - \frac{M}{\sigma L_r}\omega\psi_a
\end{aligned}
$$



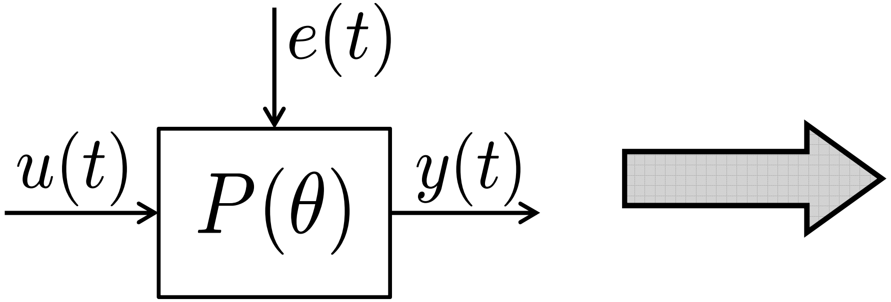
Induction Motor Scheme

Input: u_a, u_b, stator voltages

Output: i_a, i_b, stator currents

Exogenous disturbance: load torque T_l

The parameter estimation problem

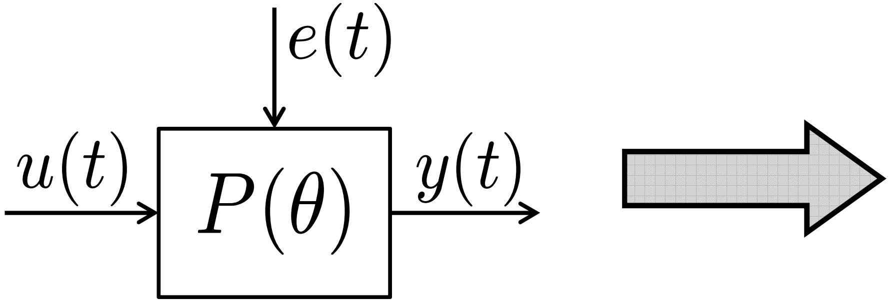


$$\left\{ \begin{array}{c} u(1), u(2), \ldots, u(N) \\ y(1), y(2), \ldots, y(N) \end{array} \right\}$$

$$\begin{aligned} \widehat{f} \quad : \quad & \mathbb{R}^{2N} \to \mathbb{R}^{q} \\ & \{u(1), y(1), \ldots, u(N), y(N)\} \to \widehat{\theta} \end{aligned}$$

The estimator is the map between available data and the parameter estimate

The parameter estimation problem



$$\left\{ \begin{array}{c} u(1), u(2), \ldots, u(N) \\ y(1), y(2), \ldots, y(N) \end{array} \right\}$$

$$\begin{aligned} \widehat{f} \quad : \quad & \mathbb{R}^{2N} \to \mathbb{R}^q \\ & \{u(1), y(1), \ldots, u(N), y(N)\} \to \widehat{\theta} \end{aligned}$$

The estimator is the map between available data and the parameter estimate

Estimator must be evaluated based on:

1. Consistency: estimate close enough to the actual parameter value for all values of θ in admissible set
2. Computational effort: time required for returning the parameter estimate

3) Available estimation tools

3.1) Prediction error methods

3.2) Kalman filtering methods

System id tools

- PE
  - Prediction error methods (LS, ML, …)

System id tools: PE

- PE - Prediction error methods:

  a) data
  b) class of models
  c) predictive criterion

- basic idea: a model is good if prediction error small

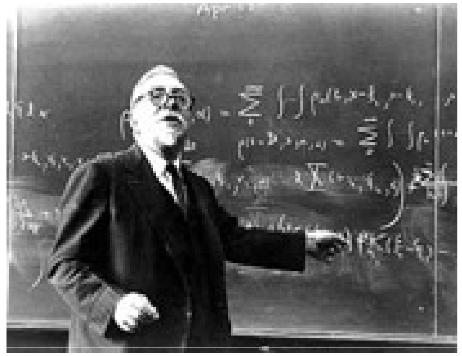
- widely used for black-box modeling, but also for parameter estimation in white-box models.

System id tools: PE

- PE - Prediction Error methods
  - Back to the origins

System id tools: PE

- 1940
  - Andrej Kolmogorov
  - Norbert Wiener



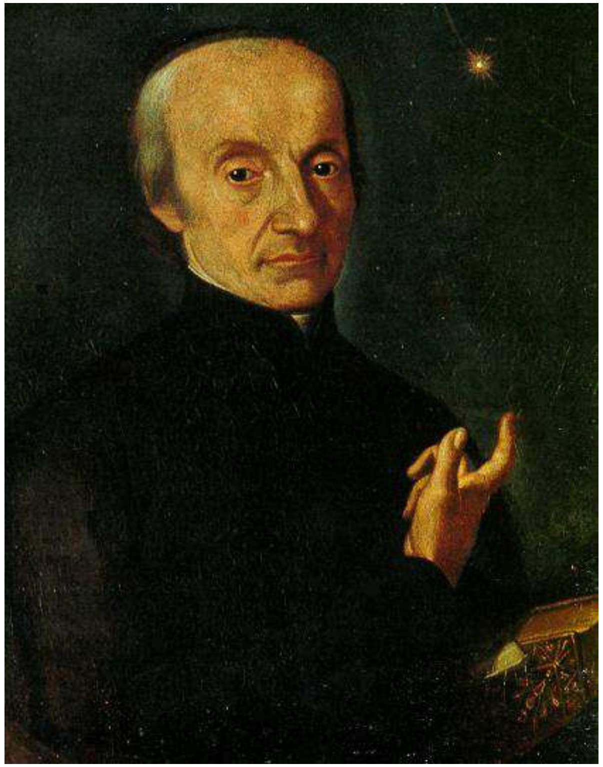
- 1605
  - Galileo Galilei:
    la pietra lavagna è la pietra
    del paragone delli ingegni
    (the blackboard is
    the test bed of all skills)

System id tools: PE

- PE - Prediction Error methods
  - Back again to the origins

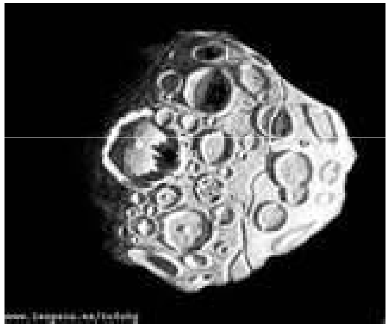
System id tools: PE

- PE - Prediction Error methods
  - 1801, Giuseppe Piazzi, astronomer in Palermo (Italy)



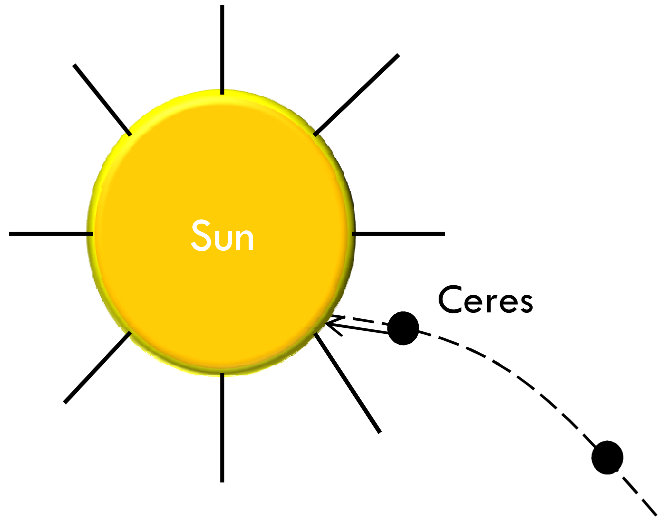
System id tools: PE

- PE - Prediction Error methods
  - 1 January 1801
    Piazzi sees a new celestial body
    and calls it Ceres
  - asteroid (dwarf planet)
    located in the position of the
    “missing planet”
    between Mars and Jupiter
  - popular emotional event



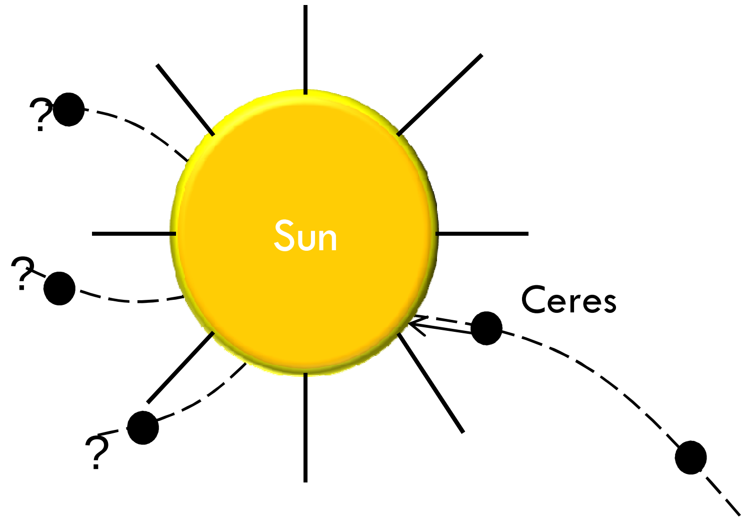
System id tools: PE

- PE - Prediction Error methods
  - after two months Ceres disappeared behind the Sun



System id tools: PE

- PE - Prediction Error methods
  - question:
    where had to re-appear ?
  - trajectory prediction



System id tools: PE

- PE – Prediction Error methods
  - trajectory Prediction
  - Carl Gauss (23) solves this prediction problem via Least Squares

nice reading
Mobius ands his band:
Mathematics and Astronomy in Nineteenth-Century Germany
John Fauvel (ed)

System id tools: PE

- PE - Prediction Error methods
  - Least Squares

    ⇨

    concept of prediction error enters the realm of science

System id tools: PE

- PE - Prediction Error methods
  - Least Squares

    ⇨

    concept of prediction error enters the realm of science

    ⇨

    all models are wrong

System id tools: PE

- PE - Prediction Error methods
  - Least Squares

    ⇨

    concept of prediction error enters the realm of science

    ⇨

    all models are wrong
    some are useful

Test-bed problem

$$
\begin{aligned}
x_1(t+1) &= \frac{1}{2}x_1(t) + u(t) + v_{11}(t) \\
x_2(t+1) &= (1-\theta^2)\sin(50\,\theta^2)\cdot x_1(t) - \theta\cdot x_2(t) + \frac{\theta}{1+\theta^2}\cdot u(t) + v_{12}(t) \\
y(t) &= x_2(t) + v_2(t)
\end{aligned}
$$

$u(t) \sim WGN(0,1)$

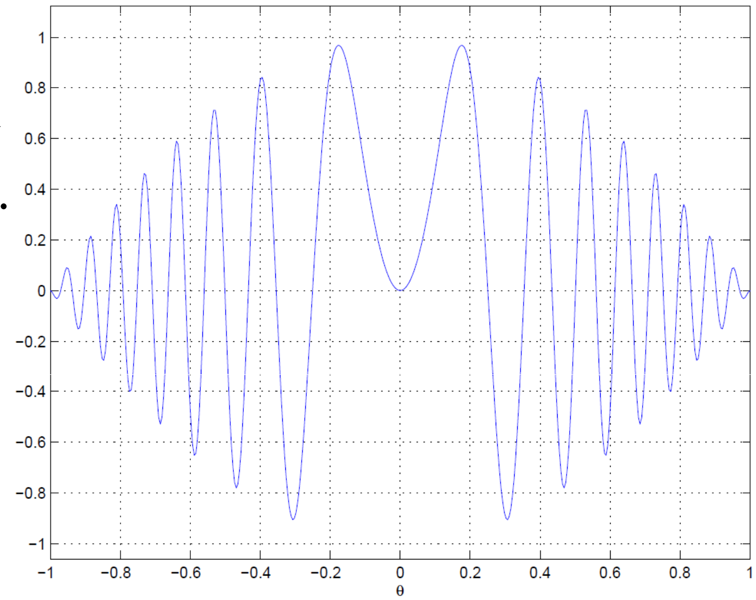
θ = uncertain parameter in [-0.9,0.9]

Test-bed problem

$$x_1(t+1) = \frac{1}{2}x_1(t) + u(t) + v_{11}(t)$$

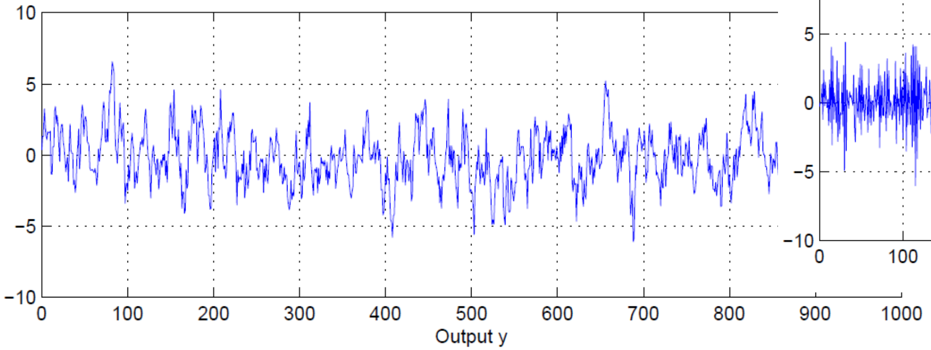
$$x_2(t+1) = (1-\theta^2)\sin(50\,\theta^2) \cdot x_1(t) - \theta \cdot$$

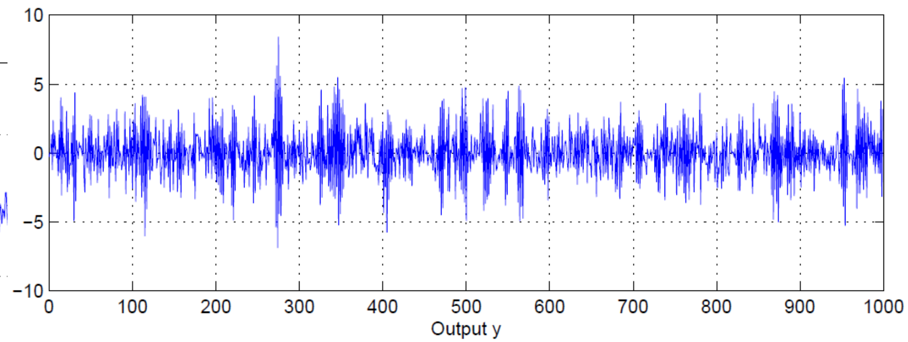
$$y(t) = x_2(t) + v_2(t)$$



$u(t) \sim WGN(0,1)$

θ = uncertain parameter in [-0.9,0.9]





Test-bed problem

$$
\begin{aligned}
x_1(t+1) &= \frac{1}{2}x_1(t) + u(t) + v_{11}(t) \\
x_2(t+1) &= (1-\theta^2)\sin(50\,\theta^2)\cdot x_1(t) - \theta\cdot x_2(t) + \frac{\theta}{1+\theta^2}\cdot u(t) + v_{12}(t) \\
y(t) &= x_2(t) + v_2(t)
\end{aligned}
$$

$u(t) \sim WGN(0,1)$

θ = uncertain parameter in [-0.9,0.9]

Data: 200 experiments:
200 values for θ
200 data sequences of length N=1000 for each θ

Test-bed problem

$$
\begin{aligned}
x_1(t+1) &= \frac{1}{2}x_1(t) + u(t) + v_{11}(t) \\
x_2(t+1) &= (1-\theta^2)\sin(50\,\theta^2)\cdot x_1(t) - \theta\cdot x_2(t) + \frac{\theta}{1+\theta^2}\cdot u(t) + v_{12}(t) \\
y(t) &= x_2(t) + v_2(t)
\end{aligned}
$$

$u(t) \sim WGN(0,1)$

θ = uncertain parameter in [-0.9,0.9]

Data: 200 experiments:
200 values for θ
200 data sequences of length N=1000 for each θ

Objective: $\left\{ \begin{array}{c} u(1), u(2), \ldots, u(N) \\ y(1), y(2), \ldots, y(N) \end{array} \right\} \longrightarrow \widehat{\theta}$

Prediction Error methods

Output predictor based on the plant model: $\widehat{y}(i, \theta)$

$$\widehat{\theta} = \arg\min \sum_{i=1}^{N} \left(y(i) - \widehat{y}(i, \theta)\right)^2$$

Estimator implicitly defined by the minimization process

Prediction Error Methods

Output predictor based on the plant model: $\widehat{y}(i,\theta)$

$$\widehat{\theta} = \arg\min \sum_{i=1}^{N} \left(y(i) - \widehat{y}(i,\theta)\right)^2$$

Estimator implicitly defined by the minimization process

Drawbacks:

1. Computation of gradient not easy (nonlinear, infinite dimensional model)

Prediction Error Methods

Output predictor based on the plant model: $\widehat{y}(i, \theta)$

$$\widehat{\theta} = \arg\min \sum_{i=1}^{N} \left(y(i) - \widehat{y}(i, \theta)\right)^2$$

Estimator implicitly defined by the minimization process

Drawbacks:

1. Computation of gradient not easy (nonlinear, infinite dimensional model)
2. Minimization is nontrivial

Prediction Error Methods

Output predictor based on the plant model: $\widehat{y}(i,\theta)$

$$\widehat{\theta} = \arg\min \sum_{i=1}^{N} \left(y(i) - \widehat{y}(i,\theta)\right)^2$$

Estimator implicitly defined by the minimization process

Drawbacks:

1. Computation of gradient not easy (nonlinear, infinite dimensional model)
2. Minimization is nontrivial

local minima – multiple initializations

Prediction Error Methods

Output predictor based on the plant model: $\widehat{y}(i,\theta)$

$$\widehat{\theta} = \arg\min \sum_{i=1}^{N} \left(y(i) - \widehat{y}(i,\theta)\right)^2$$

Estimator implicitly defined by the minimization process

Drawbacks:

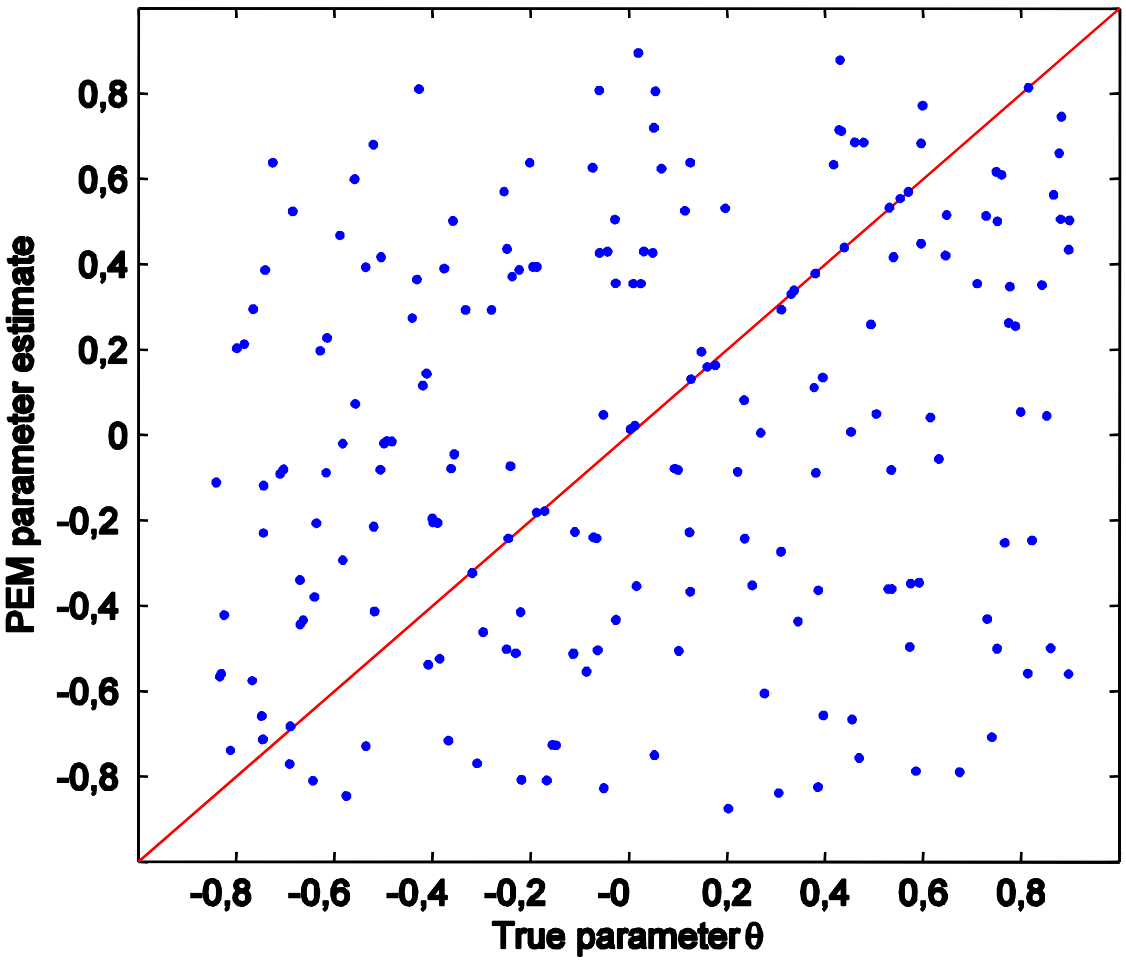
1. Computation of gradient not easy (nonlinear, infinite dimensional model)
2. Minimization is nontrivial

local minima – multiple initializations

⇩

3. High computational burden required to work out $\widehat{\theta}$

Test-bed problem – PE



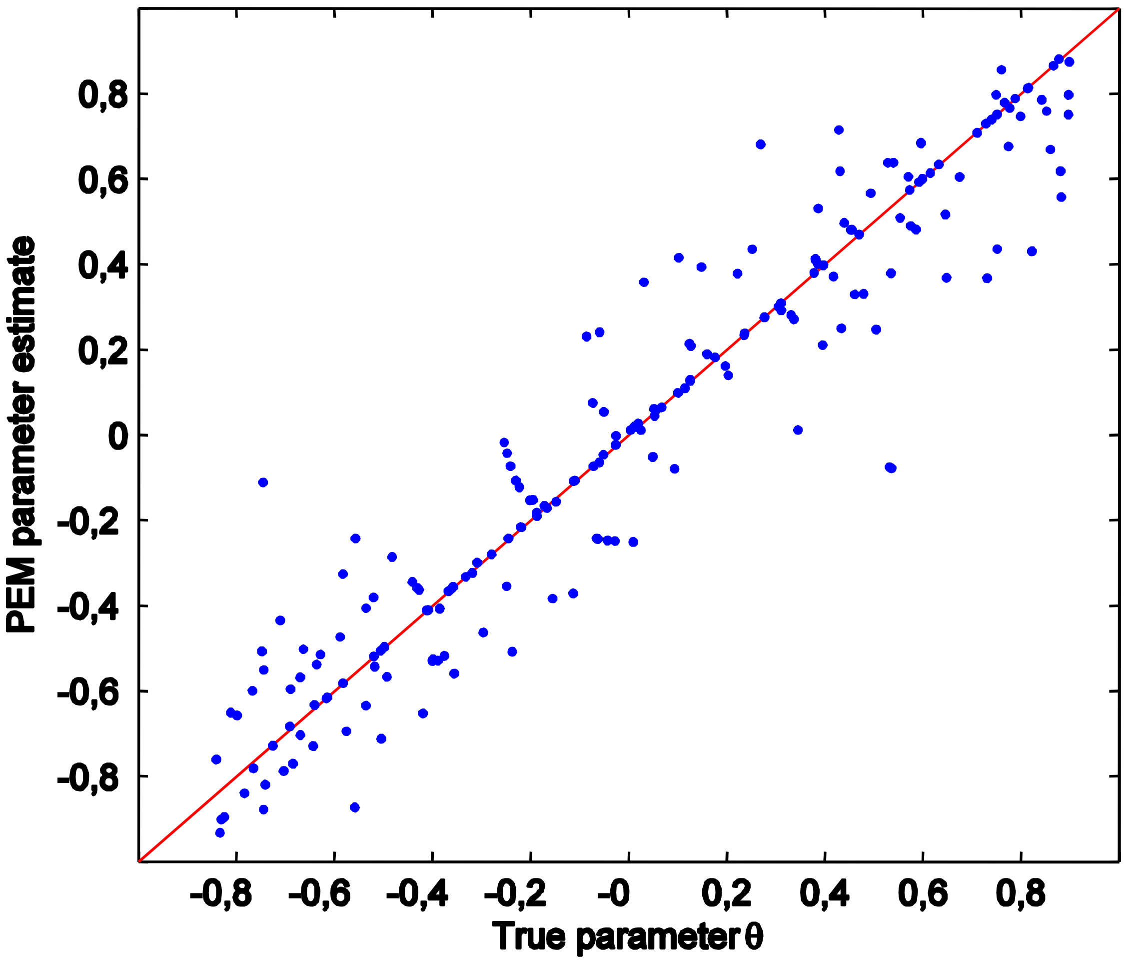
1 initialization at random

Computational time:

18.23 seconds (0.09 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – PE



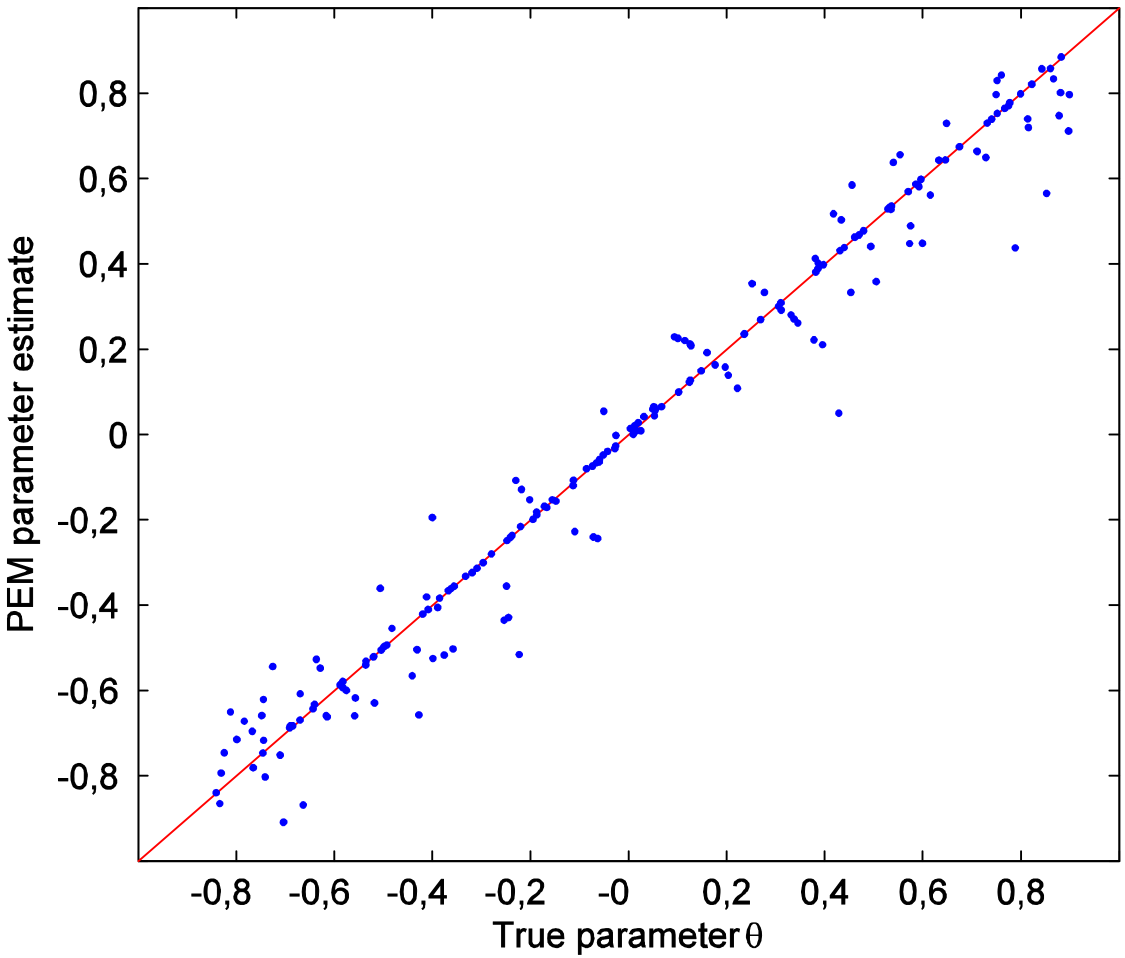
5 initializations at random

Computational time:

94.28 seconds (0.47 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – PE



10 initializations at random

Computational time:

179.14 seconds (0.89 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

System id tools: KF

- 1960
  - Rudolph R. Kalman
  - State space approach: estimate the state from input-output data

Kalman Filtering Methods

- *State space models*
- unknown parameter transformed into a state variable:

$$\theta(t+1) = \theta(t)$$

- non-linear realm

Kalman Filtering Methods

- unknown parameter transformed into a state variable:

$$\theta(t+1) = \theta(t)$$

- non-linear realm

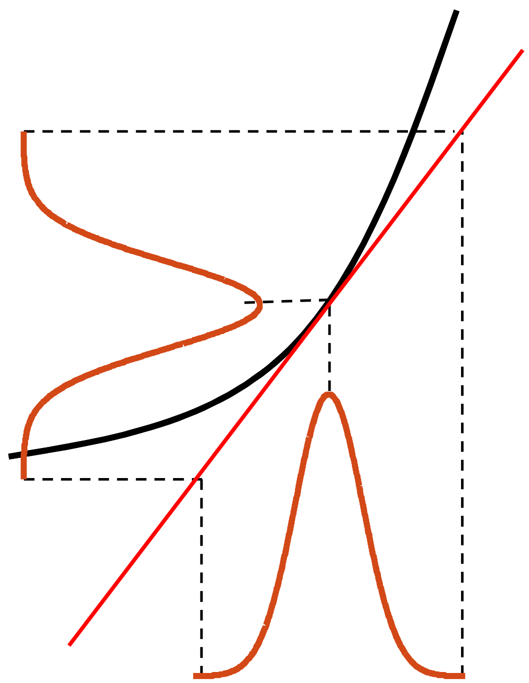
EKF
Extended Kalman Filter

PKF
Particle Kalman Filter

UKF
Unscented Kalman Filter

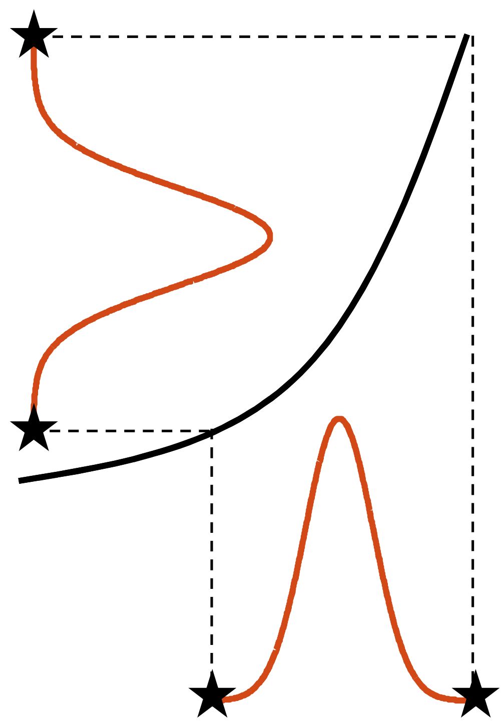
Kalman Filtering Methods

- *EKF*
  - System linearization around the last obtained state estimate
- *UKF*
  - A few representative particles (σ-points).
- *PKF*
  - A whole set of particles representative of the whole state distribution



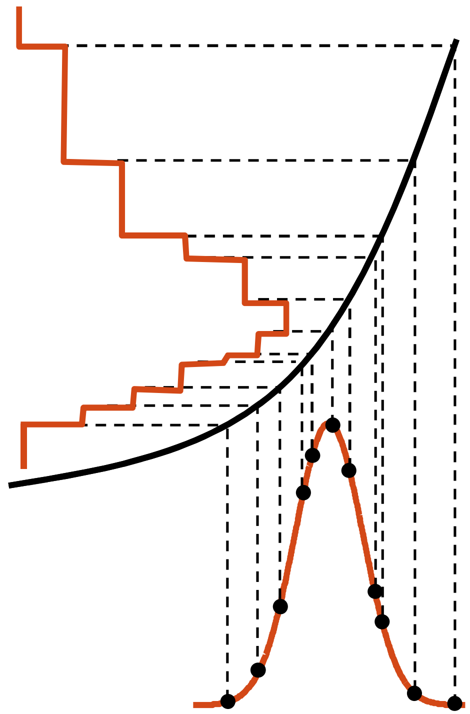
Kalman Filtering Methods

- *EKF*
  - System linearization around the last obtained state estimate
- *UKF*
  - A few representative particles (σ-points).
- *PKF*
  - A whole set of particles representative of the whole state distribution



Kalman Filtering Methods

- *EKF*
  - System linearization around the last obtained state estimate
- *UKF*
  - A few representative particles (σ-points).
- *PKF*
  - A whole set of particles representative of the whole state distribution



Kalman Filtering Methods

Additional state variable: $\theta(t+1) = \theta(t)$

$$\widehat{\theta}(t|t) = \mathbf{KF}(u(1), y(1), \ldots, u(t), y(t))$$

$$\widehat{\theta} = \widehat{\theta}(N|N)$$

EKF, UKF, PKF

$\widehat{\theta}$

Kalman Filtering Methods

Additional state variable: $\theta(t+1) = \theta(t)$

$$\widehat{\theta}(t|t) = \mathbf{KF}(u(1), y(1), \ldots, u(t), y(t))$$

$$\widehat{\theta} = \widehat{\theta}(N|N)$$

EKF, UKF, PKF

Remark:

In practice the parameter equation is: $\vartheta(t+1) = \vartheta(t) + (\, fake\, noise\,)$

question: variance of fake noise ?

question: variance of $\vartheta(0)$?

Kalman Filtering Methods

Additional state variable: $\theta(t+1) = \theta(t)$

$$\widehat{\theta}(t|t) = \mathbf{KF}(u(1), y(1), \ldots, u(t), y(t))$$

$$\widehat{\theta} = \widehat{\theta}(N|N)$$

EKF, UKF, PKF

Drawbacks:

1. Derivation of the filter equations may be not easy (nonlinear, infinite dimensional model)
2. Serious convergence problems (EKF & UKF)
3. High computational burden required to work out $\widehat{\theta}$ (PKF)

Test-bed problem – (cont'd)

$$\begin{aligned} x_1(t+1) &= \frac{1}{2}x_1(t) + u(t) + v_{11}(t) \\ x_2(t+1) &= (1-\theta^2)\sin(50\,\theta^2)\cdot x_1(t) - \theta \cdot x_2(t) + \frac{\theta}{1+\theta^2}\cdot u(t) + v_{12}(t) \\ y(t) &= x_2(t) + v_2(t) \end{aligned}$$

$u(t) \sim WGN(0,1)$

θ = uncertain parameter in [-0.9,0.9]

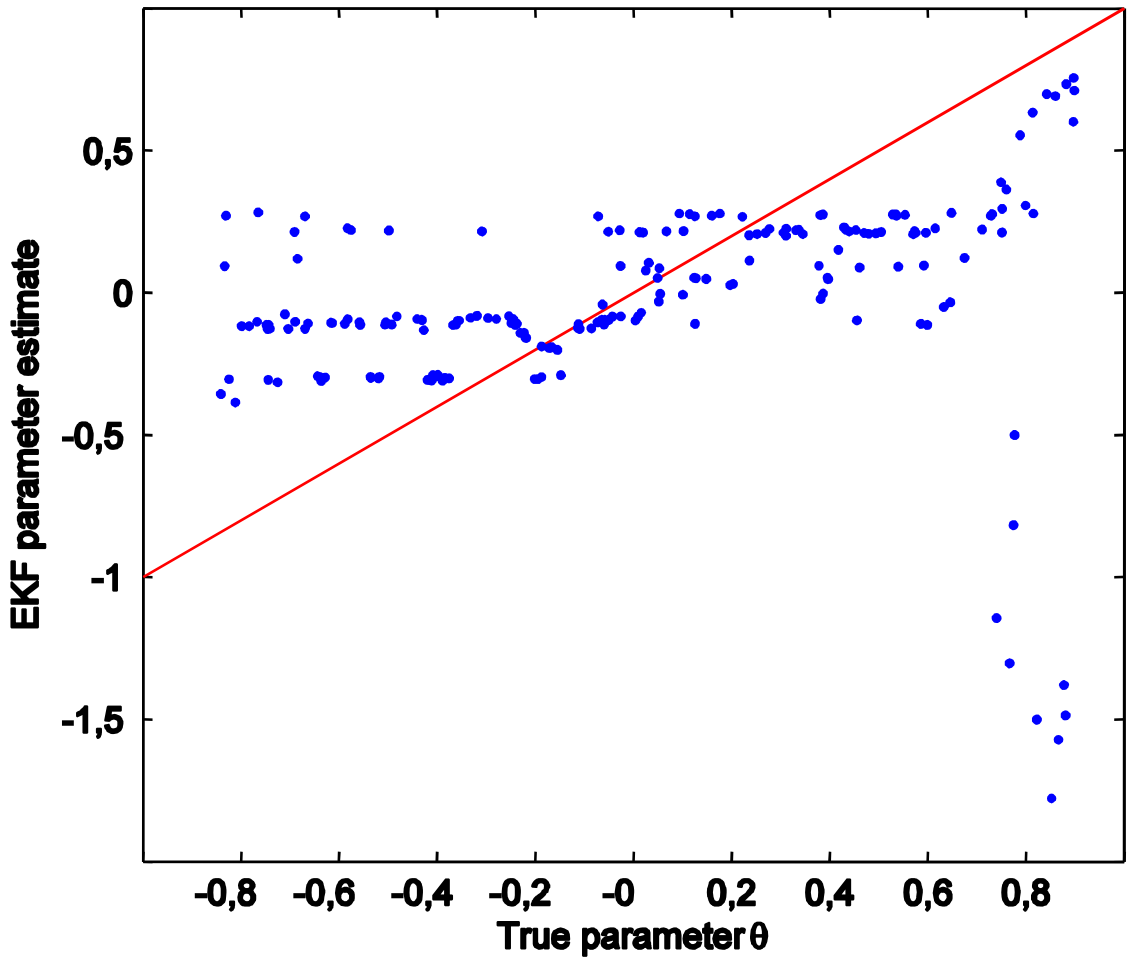
Data: 200 experiments:

200 values for θ

200 data sequences of length N=1000 for each θ

Objective: $\left\{ \begin{array}{c} u(1), u(2), \ldots, u(N) \\ y(1), y(2), \ldots, y(N) \end{array} \right\} \longrightarrow \widehat{\theta}$

Test-bed problem – EKF



Initialization

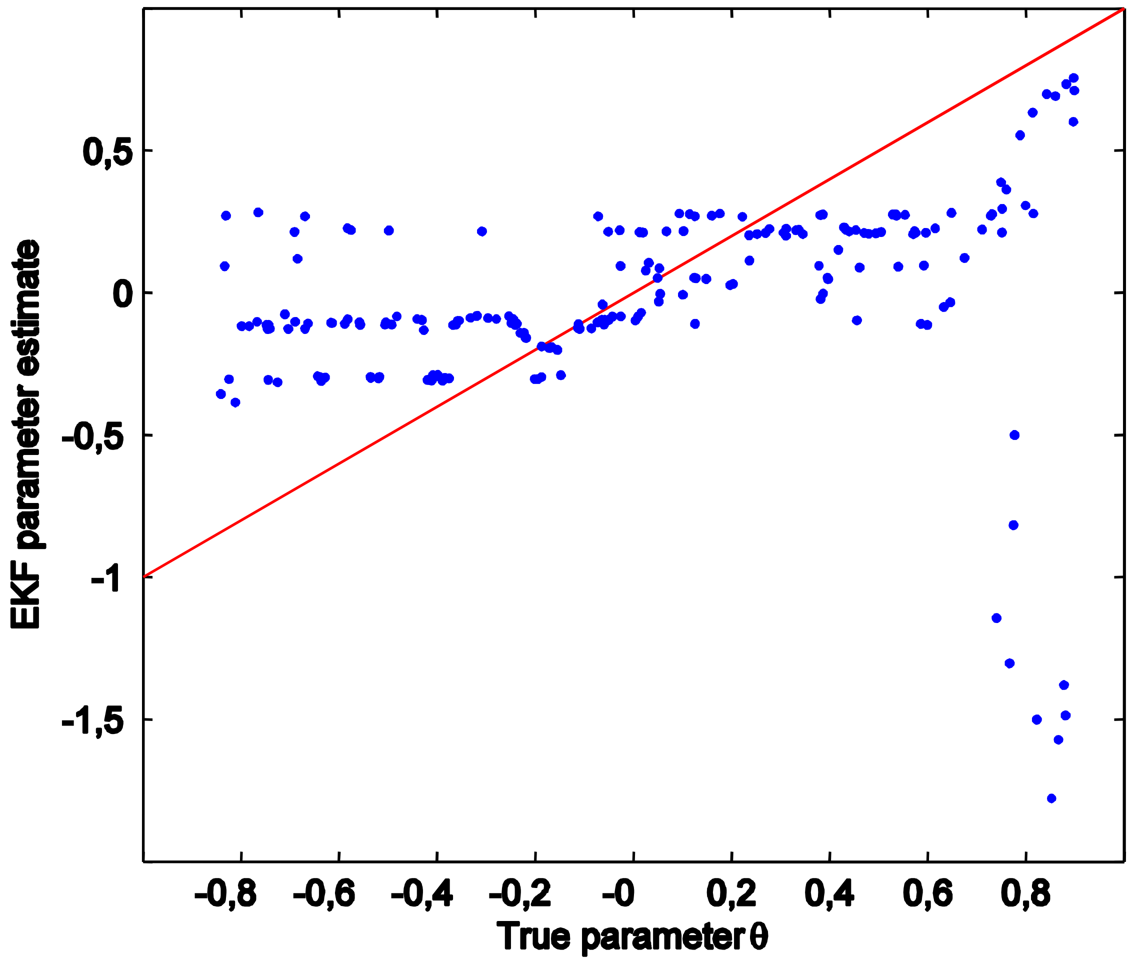
$$P_0 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 10^{-2} \end{bmatrix}$$

Computational time:

11.01 seconds (0.055 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – EKF



Initialization

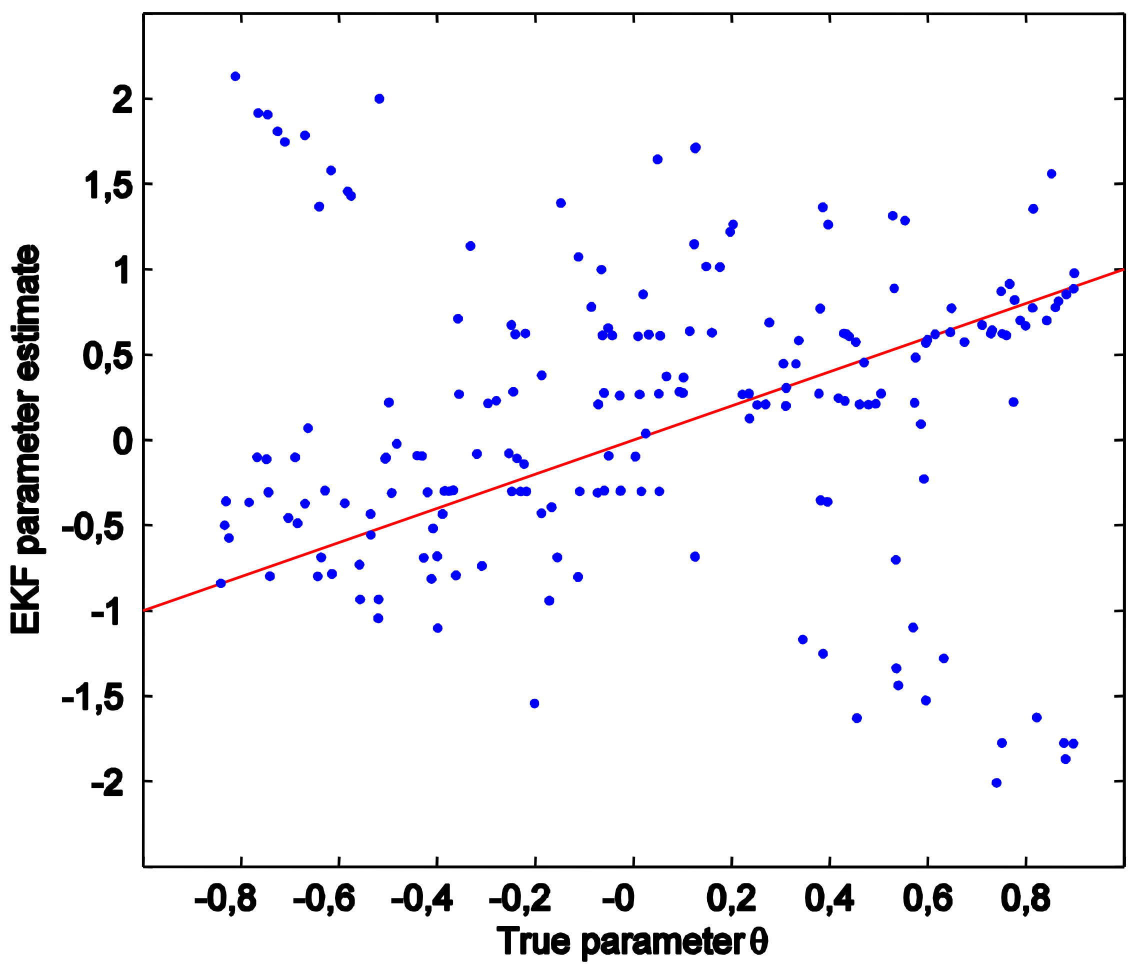
$$P_0 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 10^{-2} \end{bmatrix}$$

Computational time:

11.01 seconds
(0.055 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – EKF



Initialization

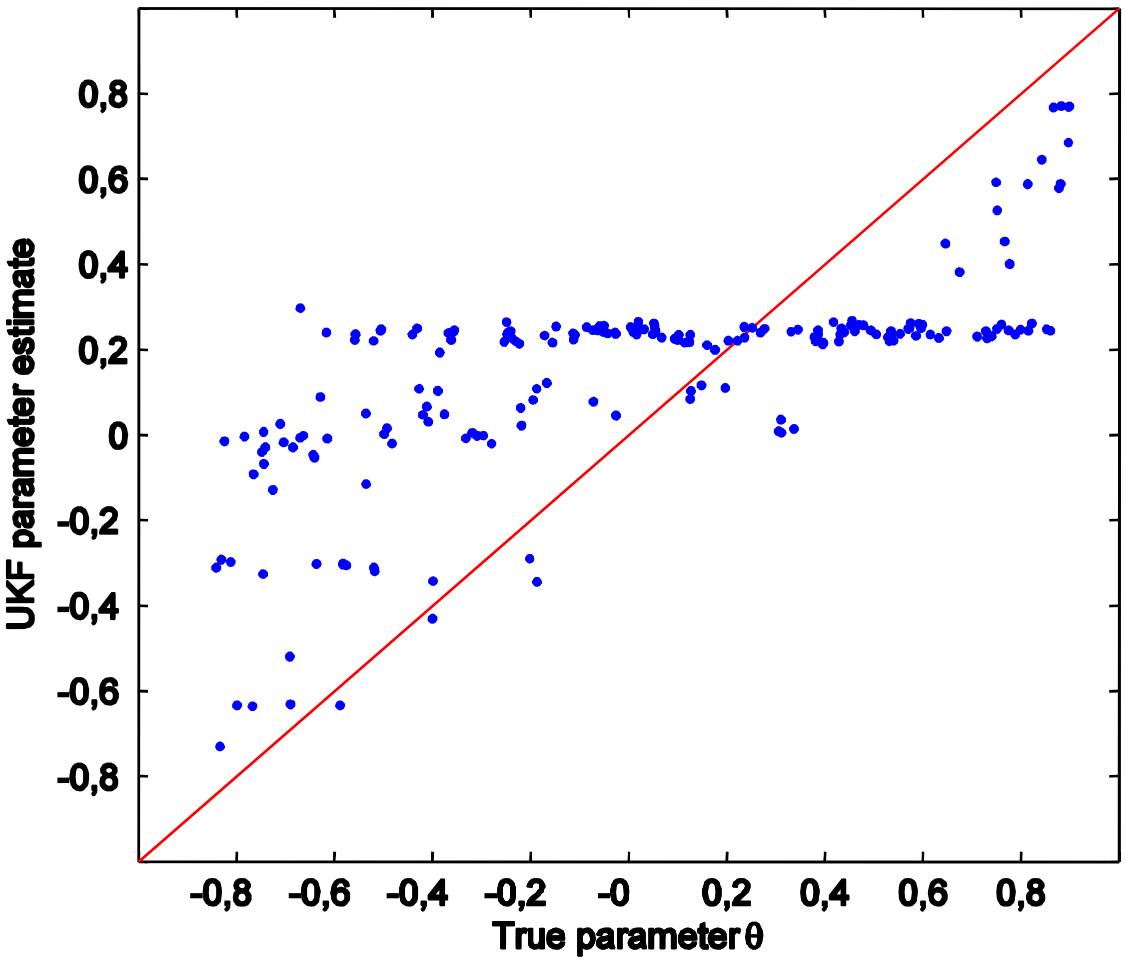
$$P_0 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

Computational time:

11.01 seconds (0.055 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – UKF



Initialization

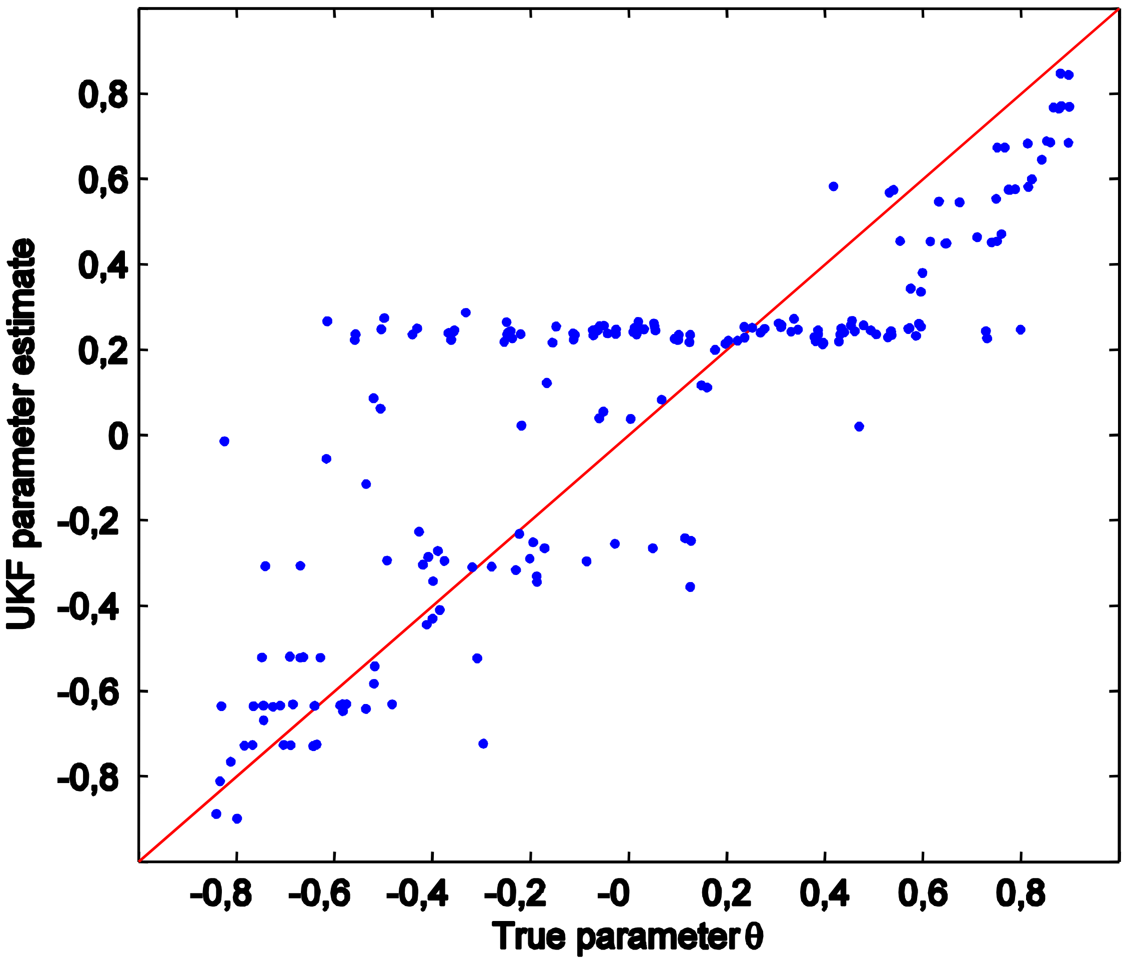
$$P_0 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 10^{-2} \end{bmatrix}$$

Computational time:

200 seconds
(1 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – UKF



Initialization

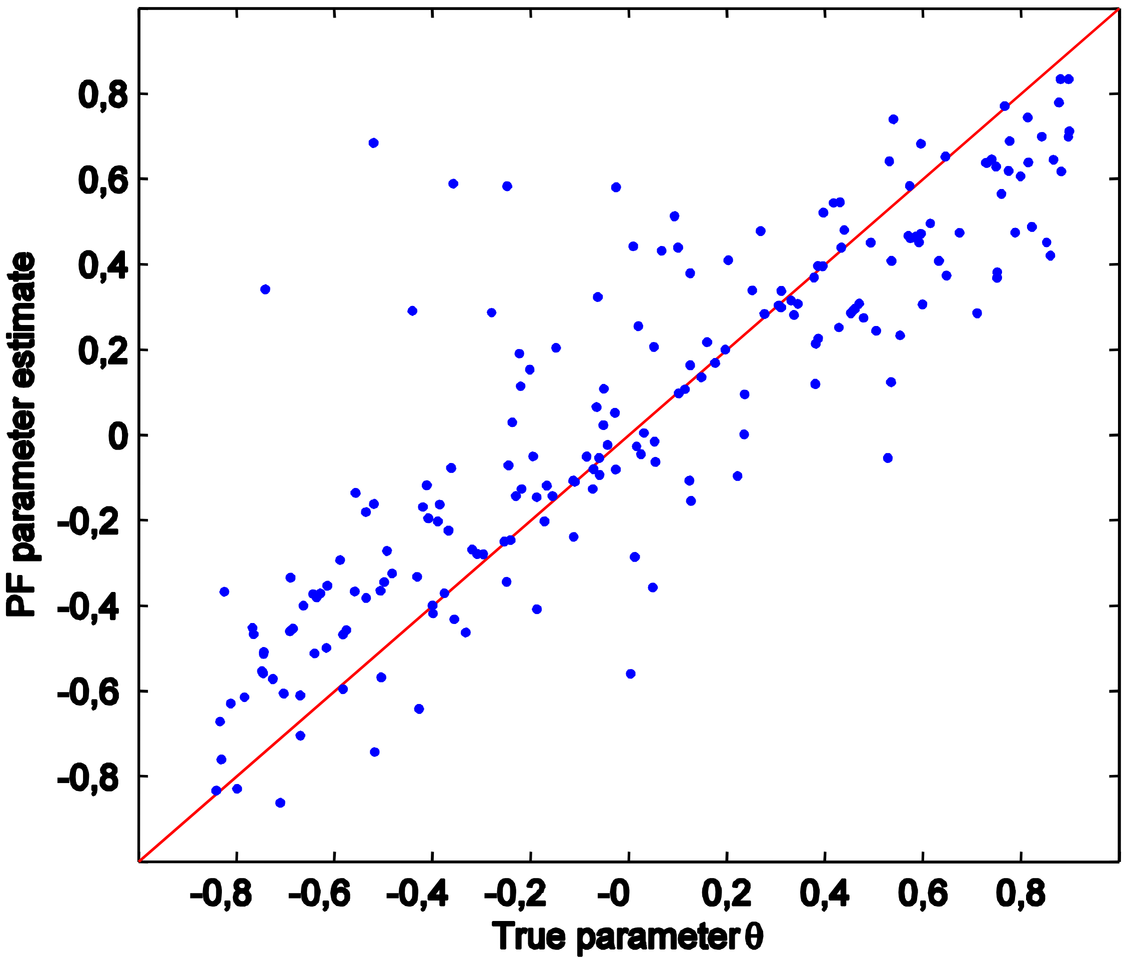
$$P_0 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

Computational time:

200 seconds
(1 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

Test-bed problem – PF



1000 particles

Computational time:

4675 seconds
(23.38 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

4) Proposing a new paradigm: the two stage (TS) method

The two stage (TS) method

with Simone Garatti

The two-stage (TS) approach

Idea: perform intensive simulation trials
from which reconstruct an explicit expression for the estimator

θ^1	$\{y^1(1), u^1(1), \ldots, y^1(N), u^1(N)\}$
θ^2	$\{y^2(1), u^2(1), \ldots, y^2(N), u^2(N)\}$
$\vdots$	$\vdots$
θ^m	$\{y^m(1), u^m(1), \ldots, y^m(N), u^m(N)\}$

Simulated
data chart

The two-stage (TS) approach

Idea: perform intensive simulation trials
from which reconstruct an explicit expression for the estimator

θ^1	$\{y^1(1), u^1(1), \ldots, y^1(N), u^1(N)\}$
θ^2	$\{y^2(1), u^2(1), \ldots, y^2(N), u^2(N)\}$
$\vdots$	$\vdots$
θ^m	$\{y^m(1), u^m(1), \ldots, y^m(N), u^m(N)\}$

Simulated
data chart

Goal: reconstruct from the chart the relationship between data and parameter:

$$\widehat{f} \leftarrow \min_{f \in \mathcal{F}} \frac{1}{m} \sum_{i=1}^{m} \left\| \theta^i - f(y^i(1), u^i(1), \ldots, y^i(N), u^i(N)) \right\|^2$$

The two-stage (TS) approach

Idea: perform intensive simulation trials
from which reconstruct an explicit expression for the estimator

θ^1	$\{y^1(1), u^1(1), \ldots, y^1(N), u^1(N)\}$
θ^2	$\{y^2(1), u^2(1), \ldots, y^2(N), u^2(N)\}$
$\vdots$	$\vdots$
θ^m	$\{y^m(1), u^m(1), \ldots, y^m(N), u^m(N)\}$

Simulated
data chart

Reconstruct from the chart the relationship between data and θ^i :

$$\widehat{f} \leftarrow \min_{f \in \mathcal{F}} \frac{1}{m} \sum_{i=1}^{m} \left\| \theta^i - f(y^i(1), u^i(1), \ldots, y^i(N), u^i(N)) \right\|^2$$

Choice of the class of functions $\mathcal{F}$ is critical:

- (bias vs. variance issue – $\widehat{f} : \mathbb{R}^{2N} \to \mathbb{R}^q$)

The two-stage (TS) approach

Idea: perform intensive simulation trials
from which reconstruct an explicit expression for the estimator

θ^1	$\{y^1(1), u^1(1), \ldots, y^1(N), u^1(N)\}$
θ^2	$\{y^2(1), u^2(1), \ldots, y^2(N), u^2(N)\}$
$\vdots$	$\vdots$
θ^m	$\{y^m(1), u^m(1), \ldots, y^m(N), u^m(N)\}$

Simulated
data chart

Reconstruct from the chart the relationship between data and θ^i :

$$\widehat{f} \leftarrow \min_{f \in \mathcal{F}} \frac{1}{m} \sum_{i=1}^{m} \left\| \theta^i - f(y^i(1), u^i(1), \ldots, y^i \;\cdots\; {}^i(N)) \right\|^2$$

MISSION
IMPOSSIBLE

Choice of the class of functions $\mathcal{F}$ is critical:

- (bias vs. variance issue – $\widehat{f} : \mathbb{R}^{2N} \rightarrow \mathbb{R}^q$)

TS Step 1: data compression

simple idea:

I/O data sequence compactly described by an ARX model:

$$
\begin{aligned}
y^i(t) \quad &= \quad \alpha_1^i y^i(t-1) + \cdots \alpha_{n_y}^i y^i(t-n_y) + \\
&\quad\ \alpha_{n_y+1}^i u^i(t-1) + \cdots + \alpha_{n_y+n_u}^i u^i(t-n_u),
\end{aligned}
$$

$$\boxed{n = n_y + n_u \ll 2N}$$

Parameter found via least squares:

$$\varphi^i(t) = [y^i(t-1) \cdots y^i(t-n_y)\ u^i(t-1) \cdots u^i(t-n_u)]^T$$

$$\begin{bmatrix} \alpha_1^i \\ \vdots \\ \alpha_n^i \end{bmatrix} = \Big[\sum_{t=1}^{N} \varphi^i(t)\varphi^i(t)^T\Big]^{-1} \cdot \sum_{t=1}^{N} \varphi^i(t) y^i(t)$$

TS Step 1: data compression

simple idea: repeat compression for all the *m* simulation trials:

θ^1	$\{\alpha_1^1, \ldots, \alpha_n^1\}$
θ^2	$\{\alpha_1^2, \ldots, \alpha_n^2\}$
$\vdots$	$\vdots$
θ^m	$\{\alpha_1^m, \ldots, \alpha_n^m\}$

Compressed artificial data chart

$n \ll 2N$, problem dimensionality reduced

Remark

$\{\alpha_1^i, \ldots, \alpha_n^i\}$ have no physical meaning

They play a purely instrumental role for reducing the size of the complexity

TS Step 2: par. Est. from compressed data

simple idea: repeat this process for all the *m* simulation trials:

θ^1	$\{\alpha_1^1, \ldots, \alpha_n^1\}$
θ^2	$\{\alpha_1^2, \ldots, \alpha_n^2\}$
$\vdots$	$\vdots$
θ^m	$\{\alpha_1^m, \ldots, \alpha_n^m\}$

Compressed artificial data chart

$n \ll 2N$, problem dimensionality reduced

Reconstruct from the artificial chart the relationship between compressed data $\{\alpha_1^i, \ldots, \alpha_n^i\}$ and θ^i:

$$\widehat{h} \leftarrow \min_{h \in \mathcal{H}} \frac{1}{m} \sum_{i=1}^{m} \left\| \theta^i - h(\alpha_1^i, \ldots, \alpha_n^i) \right\|^2$$

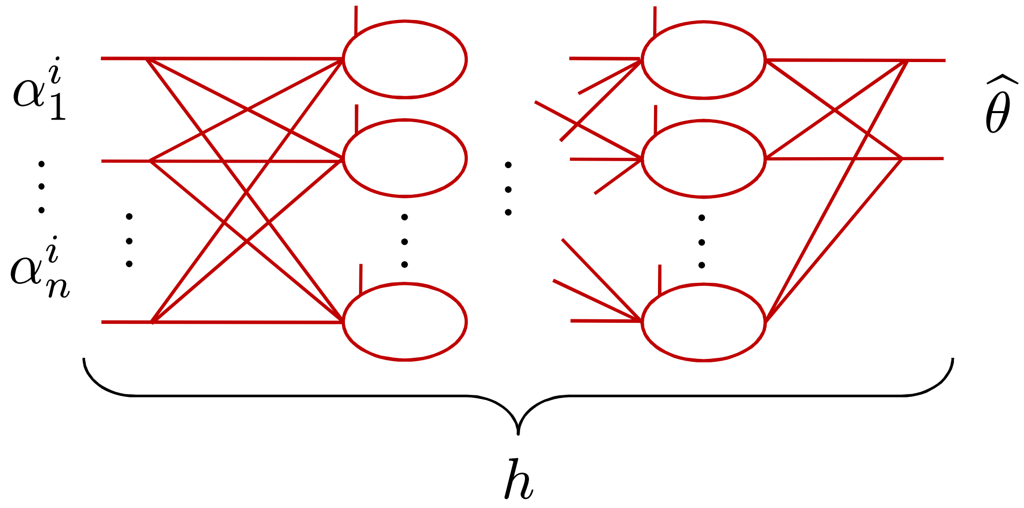
Choice of $\mathcal{H}$ is no more critical: $\widehat{h} : \mathbb{R}^n \rightarrow \mathbb{R}^q$

TS Step 2: choice of $\mathcal{H}$

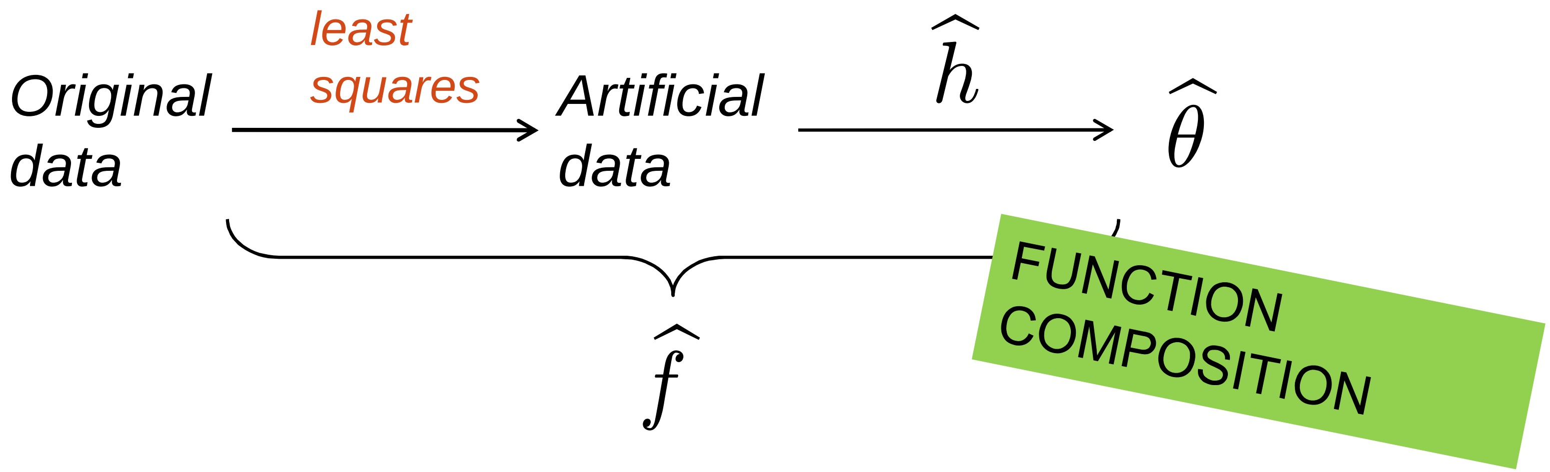
Linear regressions of $\{\alpha_1^i, \ldots, \alpha_n^i\}$

$$h = \begin{bmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & \ddots & \vdots \\ c_{q,1} & \cdots & c_{q,n} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1^i \\ \vdots \\ \alpha_n^i \end{bmatrix}$$

Neural Networks



The whole TS estimator



$$\widehat{f}: \ \{u(1), y(1), \ldots, u(N), y(N)\} \rightarrow \widehat{\theta}$$

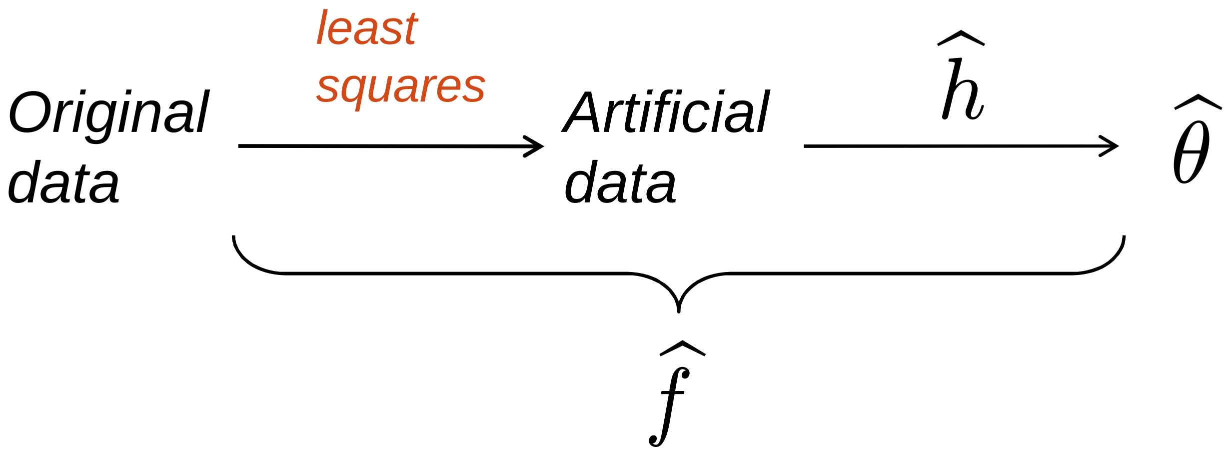
$$\widehat{f} = \widehat{h}\left(\Big[\sum_{t=1}^{N} \varphi(t)\varphi(t)^T \Big]^{-1} \cdot \sum_{t=1}^{N} \varphi(t) y(t) \right)$$

$$\varphi(t) = \begin{bmatrix} y(t-1) \\ \ldots \\ y(t-n_y) \\ u(t-1) \\ \ldots \\ u(t-n_u) \end{bmatrix}$$

TS estimator training

- Extract many values for θ in the feasible range
- For each θ, generate I/O data sequence via simulation
- Construct simulated data chart
- Fit a simple model to each data sequence
- Construct compressed artificial data chart
- Solve opt problem: compressed artificial data sequence $\Longrightarrow \theta$
- Work out the overall mapping:
  original data sequence $\Longrightarrow$ θ
  as a composition of mappings of the two steps

TS estimator at work



- Estimation:

  once a new I/O data sequence is provided,
  the corresponding θ
  is immediately found from the composition of the two steps

Test-bed problem – (cont'd)

$$
\begin{aligned}
x_1(t+1) &= \frac{1}{2}x_1(t) + u(t) + v_{11}(t) \\
x_2(t+1) &= (1-\theta^2)\sin(50\,\theta^2)\cdot x_1(t) - \theta\cdot x_2(t) + \frac{\theta}{1+\theta^2}\cdot u(t) + v_{12}(t) \\
y(t) &= x_2(t) + v_2(t)
\end{aligned}
$$

$u(t) \sim WGN(0,1)$

θ = uncertain parameter in [-0.9,0.9]

Data: 200 experiments:
200 values for θ
200 data sequences of length *N*=1000 for each θ

Objective: $\left\{ \begin{array}{c} u(1), u(2), \ldots, u(N) \\ y(1), y(2), \ldots, y(N) \end{array} \right\} \longrightarrow \widehat{\theta}$

Test-bed problem – training TS estimator

- 1500 new values for θ extracted uniformly from the interval $[-0.9, 0.9]$
- For each θ, 1000 input/output samples obtained via simulation

} Simulated data chart

- Compressed artificial data via ARX(5,5)
- $\widehat{h}$ obtained via neural network

(10 inputs, 1 outputs, 10 neurons in the 1st layer, 1 linear neuron in the output layer)

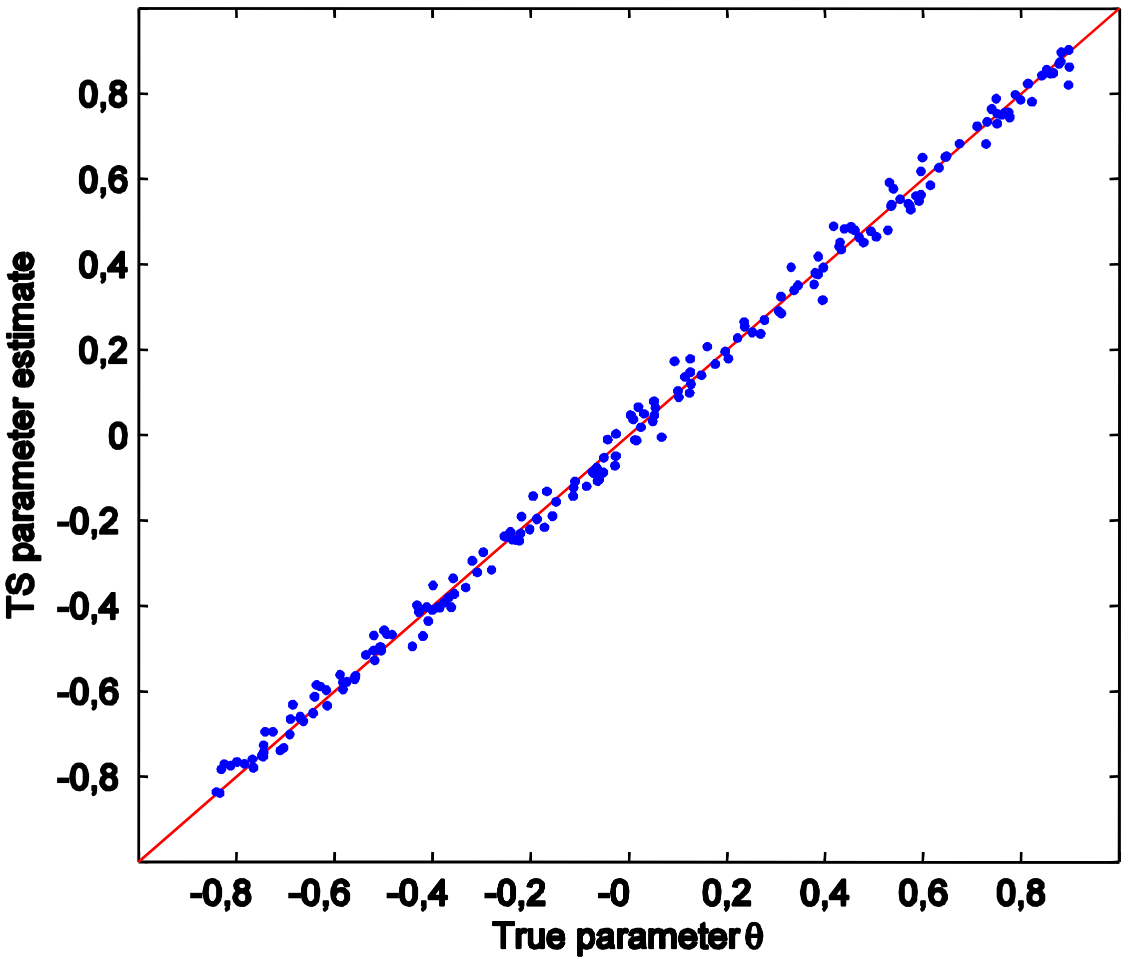
→ Models order obtained by trials (training phase is entirely off-line)

Test-bed problem – validating TS estimator

Validation of the obtained estimator over the same 200 sequences used for PEM, EKF, UKF, and PKF.

Note that the testing sequences are
different
from those used in the training phase (cross-validation)

Test-bed problem – TS



Computational time:

1.95 seconds (0.0097 seconds per estimate)

Matlab, standard 2.40 GHz dual-processor computer.

5) Conclusion

Control science 1960 ⇒ 2012

locations
and
events

System id and control 1960 ⇨ 2012

The Parameter Estimation Problem

J. H. WESTCOTT

Introduction

the variables. The method can be extended to deal with data

On the General Theory of Control Systems

R. E. KALMAN

Introduction

In no small measure, the great technological progress in automatic control and communication systems during the past two decades has depended on advances and refinements in the mathematical study of such systems. Conversely, the growth of technology brought forth many new problems (such as those related to using digital computers in control, etc.) to challenge the ingenuity and competence of research workers concerned with theoretical questions.

Despite the appearance and effective resolution of many new problems, our understanding of fundamental aspects of control has remained superficial. The only basic advance so far appears to be the *theory of information* created by Shannon[1]. The chief significance of his work in our present interpretation is the discovery of general 'laws' underlying the process of information transmission, which are quite independent of the particular models being considered or even the methods used for the description and analysis of these models. These results could be compared with the 'laws' of physics, with the crucial difference that the 'laws' governing man-made objects cannot be discovered by straightforward experimentation but only by a purely abstract analysis guided by intuition gained in observing present-day examples of technology and economic organization. We may thus classify Shannon's result as belonging to the *pure theory* of communication and control, while everything else can be labelled as the *applied theory*; this terminology reflects the well-known distinctions between pure and applied physics or mathematics. For reasons pointed out above, in its methodology the pure theory of communication and control closely resembles mathematics, rather than physics; however, it is not a branch of mathematics because at present we cannot (yet?) disregard questions of physical realizability in the study of mathematical models.

This paper initiates study of the pure theory of control, imitating the spirit of Shannon's investigations but otherwise using entirely different techniques. Our ultimate objective is to answer questions of the following type: What kind and how much information is needed to achieve a desired type of control? What intrinsic properties characterize a given unalterable plant as far as control is concerned?

At present only superficial answers are available to these questions, and even then only in special cases.

Initial results presented in this Note are far from the degree of generality of Shannon's work. By contrast, however, only *constructive* methods are employed here, giving some hope of being able to avoid the well-known difficulty of Shannon's theory: methods of proof which are impractical for actually constructing practical solutions. In fact, this paper arose from the need for a better understanding of some recently discovered computation methods of control-system synthesis[2-5]. Another by-product of the paper is a new computation method for the solution of the classical Wiener filtering problem[7].

The organization of the paper is as follows:

In Section 3 we introduce the models for which a fairly complete theory is available: dynamic systems with a finite dimensional state space and linear transition functions (i.e. systems obeying linear differential or difference equations). The class of random processes considered consists of such dynamic systems excited by an uncorrelated gaussian random process. Other assumptions, such as stationarity, discretization, single input/single output, etc., are made only to facilitate the presentation and will be absent in detailed future accounts of the theory.

In Section 4 we define the concept of *controllability* and show that this is the 'natural' generalization of the so-called 'dead-beat' control scheme discovered by Oldenbourg and Sartorius[21] and later rederived independently by Tsypkin[22] and the author[17].

We then show in Section 5 that the general problem of optimal regulation is solvable if and only if the plant is completely controllable.

In Section 6 we introduce the concept of *observability* and solve the problem of reconstructing unmeasurable state variables from the measurable ones in the minimum possible length of time.

We formalize the similarities between controllability and observability in Section 7 by means of the *Principle of Duality* and show that the Wiener filtering problem is the natural dual of the problem of optimal regulation.

Section 8 is a brief discussion of possible generalizations and currently unsolved problems of the pure theory of control.

Notation and Terminology

The reader is assumed to be familiar with elements of linear algebra, as discussed, for instance, by Halmos[8].

Consider an n-dimensional real vector space X. A *basis* in X is a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in X such that any vector $\mathbf{x}$ in X can be written uniquely as

$$\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n \qquad (1)$$

the x_i being real numbers, the *components* or *coordinates* of $\mathbf{x}$. Vectors will be denoted throughout by small bold-face letters.

The set X^* of all real-valued linear functions $\mathbf{x}^*$ (= *covectors*) on X, with the 'natural' definition of addition and scalar multiplication, is an n-dimensional vector space. The value of a covector $\mathbf{y}^*$ at any vector $\mathbf{x}$ is denoted by $[\mathbf{y}^*, \mathbf{x}]$. We call this the *inner product* of $\mathbf{y}^*$ by $\mathbf{x}$. The vector space X^* has a natural basis $\mathbf{a}^*_1, \ldots, \mathbf{a}^*_n$ associated with a given basis in X; it is defined by the requirement that

$$[\mathbf{a}^*_i, \mathbf{a}_j] = \delta_{ij} \qquad (2)$$

Using the 'orthogonality relation' 2, we may write 1 in the form

$$\mathbf{x} = \sum_{i=1}^{n} [\mathbf{a}^*_i, \mathbf{x}]\mathbf{a}_i \qquad (3)$$

which will be used frequently.

For purposes of numerical computation, a vector may be considered a matrix with one column and a covector a matrix

38 481

- System id and control

two collaborative friends

Development of id methods: IFAC SYSID's story

IFAC Symposium on System Identification (SYSID)

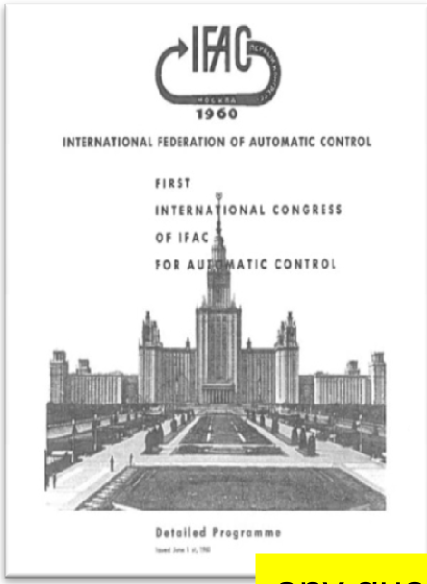
The long story of system identification congresses:

- Prague, The Check Republic (1970)
- Hague/Delft, The Netherlands (1973)
- Tblisi, URSS (1976)
- Darmstadt, Germany (1979)
- Arlington, USA (1982)
- York, UK (1985)
- Beijing, China (1988)
- Budapest, Hungary (1991)
- Copenhagen, Denmark (1994)
- Fukuoka, Japan (1997)
- Santa Barbara, USA (2000)
- Rotterdam, The Netherlands (2003)
- Newcastle, Australia (2006)
- St. Malo, France (2009)
- Brussels, Belgium (2012)
- Beijing, China (2015)

yet most interesting problems are open

Control science 1960 ⇨ 2012

- Control
- Identification



the most exiting
research area
in the planet

for 50 years
more (at least)

控制科学是在行星中最好的研究区。

any questions or comments: bittanti@elet.polimi.it