Overhead Slides for Eight Lectures on Stochastic Control and its Applications

Åström, Karl Johan

1980

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Åström, K. J. (1980). *Overhead Slides for Eight Lectures on Stochastic Control and its Applications*. (Technical Reports TFRT-7204). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

CODEN: LUTFD2/(TFRT-7204)/1-121/(1980)

OVERHEAD SLIDES FOR EIGHT LECTURES ON STOCHASTIC CONTROL AND ITS APPLICATIONS

KARL JOHAN ÅSTRÖM

DEPARTMENT OF AUTOMATIC CONTROL
LUND INSTITUTE OF TECHNOLOGY
SEPTEMBER 1980

OVERHEAD SLIDES

FOR EIGHT LECTURES ON STOCHASTIC CONTROL THEORY

AND ITS APPLICATIONS

K J Åström

LECTURES 1 & 2 - LINEAR STOCHASTIC CONTROL THEORY

The major results of linear stochastic control theory are reviewed. Emphasis is given to the problem formulation and to the main mathematical models used. The topics covered include minimum variance control, Kalman filtering, and the separation theorem. The problems are discussed both from the input-output and the state space point of view. The presentation is based on references [1], [2], and [3].

LECTURE 3 - MODELING AND SYSTEM IDENTIFICATION

The problem of obtaining the mathematical models used in linear stochastic control from experimental plant data is discussed. Principles of parameter estimation are discussed. The techniques include maximum likelihood and prediction error methods. Theoretical results for large sample properties are discussed together with computational and practical aspects. The presentation is based on references [4] and [5].

LECTURE 4 - APPLICATION OF PARAMETER ESTIMATION METHODS TO DETERMINE SHIP STEERING DYNAMICS

Application of the methods described in lecture 2 to determination of mathematical models for ship steering dynamics. The topics discussed include: mathematical models for ship dynamics, disturbances due to wind and waves, parameter estimation and practical aspects on parameter estimation. The presentation is based on references [6], [7], and [8].

LECTURE 5 - NONLINEAR STOCHASTIC CONTROL THEORY

Some important results in nonlinear stochastic control theory are discussed. The structure of the optimal feedback is derived for processes described by Markov chains using Dynamic programming. Nonlinear filtering and the nature of the solutions for more general problems are treated. Concepts of certainty equivalence, caution and probing are introduced. The presentation is based on reference [2].

LECTURES 6 & 7 - SELF-TUNING AND ADAPTIVE CONTROL

Principles of adaptive control are introduced based on nonlinear stochastic control theory. Self-tuning regulators are introduced as an approximation to the general nonlinear problem. It is also shown that self-tuning generators can be motivated in more general cases. Relations to model reference adaptive systems is discussed. Theory of adaptive control and its short-comings and applications are reviewed. The presentation is based on reference [9].

LECTURE 8 - DESIGN OF AN ADAPTIVE AUTOPILOT FOR SHIP STEERING

Describes the application of adaptive control to a practical problem. The autopilot includes gain scheduling to eliminate influence of velocity variations, a Kalman filter, and a self-tuning regulator. The design trade-offs are discussed. Results from full-scale experiments on several ships are described. The presentation is based on reference [10].

REFERENCES

[1] K J Åström: Introduction to Stochastic Control Theory, Academic Press, 1970.

[2] K J Åström: Stochastic control problems. In W A Coppel (Ed.): Mathematical Control Theory, Springer Verlag, Berlin, 1978.

[3] K J Åström: Algebraic system theory as a tool for regulator design. In Halme et al: Topics in System Theory Publication in honour of Prof Hans Blomberg on the occasion of his sixtieth birthday, Acta Polytechnica Scandinavica, Ma31, Helsinki, 1979.

[4] K J Åström: The role of identification in process modeling. VDE-Berichte Nr 276, 1977.

[5] K J Åström: Maximum likelihood and prediction error methods. Tutorial 5th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, 1979.

[6] K J Åström and C G Källström: Identification of ship-steering dynamics, Automatica 12 (1976) 9-22.

[7] C G Källström and K J Åström: Experiences of system identification applied to ship-steering dynamics. Case study 5th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, 1979.

[8] K J Åström, C G Källström, N H Norrbin, and L Byström: The identification of linear ship-steering dynamics using maximum likelihood parameter estimation. Report 1920-1, The Swedish State Shipbuilding Experimental Tank, Göteborg, Sweden, December 1974.

[9] K J Åström: Self-tuning regulators design principles and applications. To appear in K S Narendra and R V Monopoli: Applications of Adaptive Control, Academic Press.

[10] C G Källström, K J Åström, N E Thorell, J Eriksson, and L Sten: Adaptive autopilots for tankers, Automatica 15 (1979) 241-254.

STOCHASTIC CONTROL THEORY AND APPLICATIONS

1. INTRODUCTION
2. LINEAR THEORY
   - INPUT-OUTPUT MODELS
   - STATE MODELS
3. MODELING & IDENTIFICATION
4. AN APPLICATION
5. NONLINEAR THEORY
6. SELF-TUNING CONTROL
7. AN APPLICATION
8. CONCLUSIONS

STOCHASTIC CONTROL THEORY AND APPLICATIONS

1. INTRODUCTION
2. LINEAR THEORY
   - INPUT - OUTPUT MODELS
   - STATE MODELS
3. MODELING & IDENTIFICATION
4. AN APPLICATION
5. NONLINEAR THEORY
6. SELF-TUNING CONTROL
7. AN APPLICATION
8. CONCLUSIONS

NO DISTURBANCES - NO CONTROL PROBLEM

HOW TO CHARACTERIZE DISTURBANCES?

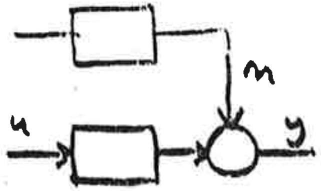
- CLASSICAL CONTROL THEORY

STEP, RAMP, SINUSOID, EXPONENTIAL
(SIGNALS GENERATED BY ODE.)

- STOCHASTIC CONTROL THEORY

STOCHASTIC PROCESSES

- MODELS:



MOSTLY LINEAR → EXTERNAL I/O

m CHARACTERIZED BY COVARIANCE FUNCTION & SPECTRAL DENSITY

CAN BE NONLINEAR → INTERNAL (STATE EQUATIONS)

$$x(t+1) = f(x(t), u(t), t) + v(x(t), u(t), t)$$

$$\sigma(x(t), u(t), t) \cdot e(t)$$

CONCEPTUAL IMPLICATION
PRACTICAL —"—

- IMPORTANT CONCEPTS

PREDICTION FILTERING SMOOTHING

EXAMPLE:

$$\frac{dy}{dt} = u \quad , \; y(0) = a$$

$$J = \int_0^\infty [y^2(t) + u^2(t)]\, dt$$

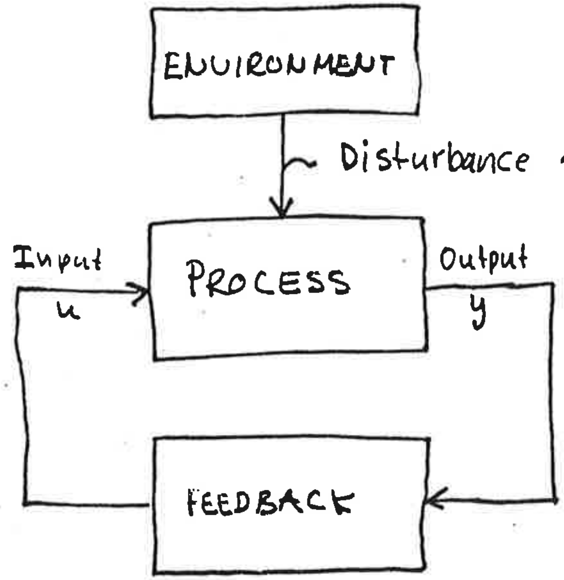
FIND u TO MINIMIZE J!

SOLUTION

(1) $u(t) = - y(0) \cdot e^{-t}$ $\qquad J = a^2$

(2) $u(t) = - y(t)$ $\qquad J = a^2$

FEEDBACK PROCESS



PI · REGULATOR

$$u(t) = u_{ref}(t) + K\left[e(t) + \int^t e(s) \frac{ds}{T}\right]$$

$$e(t) = y_{ref}(t) - y(t)$$

CONTROL ENGINEERS BLESSING

CONTROL THEORETICIANS CURSE

PROBLEM: FIND A SUITABLE THEORETICAL FRAMEWORK TO DISCUSS & ANALYSE FEEDBACK

STOCHASTIC CONTROL THEORY AND APPLICATIONS

1. INTRODUCTION
2. LINEAR THEORY
   - ☆INPUT - OUTPUT MODELS
   - STATE MODELS
3. MODELING & IDENTIFICATION
4. AN APPLICATION
5. NONLINEAR THEORY
6. SELF-TUNING CONTROL
7. AN APPLICATION
8. CONCLUSIONS

MINIMUM VARIANCE CONTROL

EXTERNAL DESCRIPTIONS

1. MATHEMATICAL MODELS

Process Dynamics
Environment — ARMAX
"CARMA" model
Multivariable generalizations

2. OPTIMAL PREDICTION

3. CONTROL

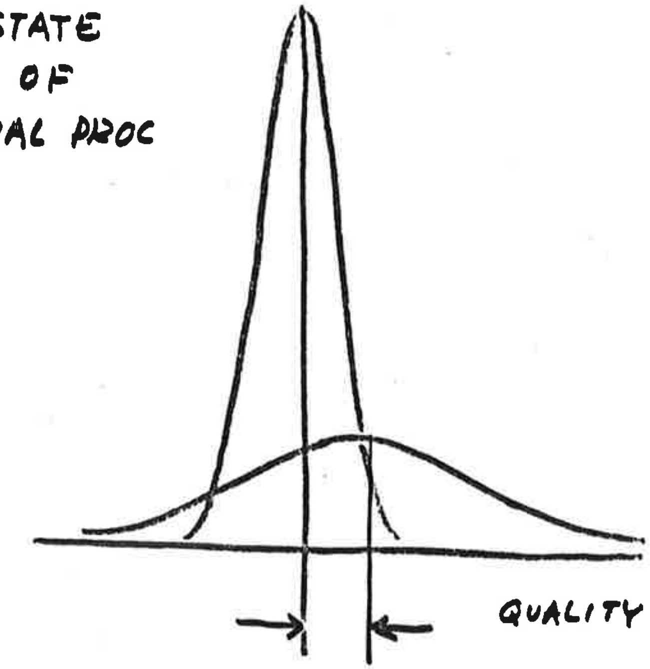
Criteria
Admissible controls
Min variance control
"Minimum Phase" condition

4. APPLICATIONS

REVIEW OF ASSUMPTIONS

- LINEAR DYNAMICS
- STOCHASTIC DISTURBANCES
- QUADRATIC LOSSFUNCTIONS

STEADY STATE
CONTROL OF
INDUSTRIAL PROC



- MODELS OF PROCESS DYNAMICS & DISTURBANCES KNOWN
  (IDENTIFICATION CAN HELP IF MODELS ARE NOT KNOWN)

- MINIMUM VARIANCE CONTROL IS EASY TO DERIVE BECAUSE ONE-STEP OPTIMIZATION IS THE SAME AS N-STEP OPTIMIZATION I:E

$$\min E \frac{1}{N} \sum_{1}^{N} y^2(t)$$

$$\min E\, y^2(t)$$

GIVES THE SAME CONTROL LAW (THIS IS OF INTEREST FOR THE ADAPTIVE PROBLEM WHERE SIMPLICITY OF CALCULATIONS IS IMPORTANT)

- ONE-STEP AND N-STEP NOT THE SAME IF THERE IS PENALTY ON CONTROL

- NOTICE

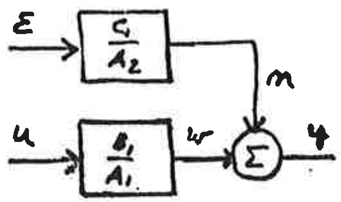
$$A y(t) = B u(t-k) + \lambda C e(t)$$

$$C = AF + q^{-k} G$$

$$u(t) = -\frac{G}{BF} y(t) \quad \text{MIN VAR CONTROL LAW}$$

CANCELLATION OF B ...

MATHEMATICAL MODELS



$$w(t) + a_1' w(t-1) + \dots + a_m w(t-m) = b_0' u(t-k) + \dots + b_m' u(t-k- \dots$$

$$A_1(q^{-1})\, w(t) = B_1(q^{-1})\, u(t-k) \qquad q\, y(t) = y(t+1)$$

$$y(t) = \frac{A_2}{A_2} \frac{B_1(q^{-1})}{A_1(q^{-1})} u(t-k) + \frac{A_1}{A_1} \frac{C_1(q^{-1})}{A_2(q^{-1})} \varepsilon(t)$$

$$A(q^{-1})\, y(t) = B(q^{-1})\, u(t-k) + C(q^{-1})\, \varepsilon(t)$$

$C(\xi)$ CAN ALWAYS BE CHOSEN TO HAVE ALL ZEROS OUTSIDE UNIT DISC OR ON UNIT CIRCLE

IF C HAS ZEROS ON THE UNIT CIRCLE THEN OPTIMAL PREDICTORS & CONTROLS ARE TIMEVARYING !

OPTIMAL PREDICTION

$$A(q^{-1})\,y(t) = C(q^{-1})\,\varepsilon(t)$$

All zeros of det $C(\xi)$ outside unit disc.

$$y(t+k) = A^{-1}(q^{-1})\,C(q^{-1})\,\varepsilon(t+k) = \mathcal{C}(q^{-1})\,\mathcal{A}^{-1}(q^{-1})\,\varepsilon(t+k)$$

$$\det A(\xi) = \det \mathcal{A}(\xi)$$

$$\mathcal{C}(q^{-1})\,\mathcal{A}^{-1}(q^{-1}) = F(q^{-1}) + q^{-k}\,\mathcal{G}(q^{-1})\,\mathcal{A}^{-1}(q^{-1})$$

F: degree $k-1$

$$y(t+k) = F(q^{-1})\,\varepsilon(t+k) + \mathcal{G}(q^{-1})\,\mathcal{A}^{-1}(q^{-1})\,\varepsilon(t)$$

$$= \underbrace{F(q^{-1})\,\varepsilon(t+k)}_{\varepsilon(t+k) + f_1\varepsilon(t+k-1) + \ldots + f_{k-1}\varepsilon(t+1)} + \mathcal{G}(q^{-1})\,\mathcal{C}^{-1}(q^{-1})\,y(t)$$

$$\tilde{y} = y(t+k) - \hat{y} = F\varepsilon(t+k) + [\mathcal{G}\mathcal{C}^{-1}y(t) - \hat{y}]$$

$$E\{a^T F\varepsilon(t+k)\}\{a^T[\mathcal{G}\mathcal{C}^{-1}y(t) - \hat{y}]\} = 0$$

Independence $\hat{y}$ arbitrary fcn of $y(t), y(t-1), \ldots$

Uncorrelated $\hat{y}$ linear in $y(t), y(t-1), \ldots$

$$\hat{y}(t+k/t) = \mathcal{G}(q^{-1})\,\mathcal{C}^{-1}(q^{-1})\,y(t)$$

$$\tilde{y}(t+k/t) = F(q^{-1})\,\varepsilon(t+k) = \varepsilon(t+k) + f_1\varepsilon(t+k-1) + \ldots + f_{k-1}\varepsilon(t+1)$$

EXAMPLE:

$$y(t+1) = \varepsilon(t+1) - \varepsilon(t)$$

$$y(t) = \varepsilon(t) - \varepsilon(t-1) \qquad C = 1 - q^{-1}$$

$$\hat{y}(t+1|t) = -\varepsilon(t)$$

$$\varepsilon(t) = \sum_{k=t_0+1}^{t} y(k) + \varepsilon(t_0) = z(t) + \varepsilon(t_0)$$

$$\hat{\varepsilon}(t_0) = -\frac{1}{t-t_0}\sum_{t_0+1}^{t} z(k) = -\frac{1}{t-t_0}\sum_{t_0+1}^{t}(t+1-k)\,y(k)$$

$$\hat{y}(t+1|t) = -\sum_{i=1}^{t-t_0}\frac{t-t_0-i}{t-t_0}\,y(t+1-i)$$

$$E(y-\hat{y})^2 \to E\varepsilon^2$$

$$\hat{\hat{y}} = -\sum y(k)$$

$$E(y-\hat{\hat{y}})^2 \to 2E\varepsilon^2$$

MINIMUM VARIANCE CONTROL

PROCESS MODEL

$$A(q^{-1})y(t) = B(q^{-1})u(t-k) + C(q^{-1})\varepsilon(t)$$

$$A(q^{-1}) = I + A_1 q^{-1} + \dots + A_n q^{-n}$$

$$B(q^{-1}) = B_0 + B_1 q^{-1} + \dots + B_n q^{-n}$$

$$C(q^{-1}) = I + C_1 q^{-1} + \dots + C_n q^{-n}$$

CRITERION

$$\min E\, y^T(t)\, Q\, y(t)$$

ASSUMPTIONS

$\det C(\xi)$ and $\det B(\xi)$ all zeros outside unit disc, $\det B_0 \neq 0$

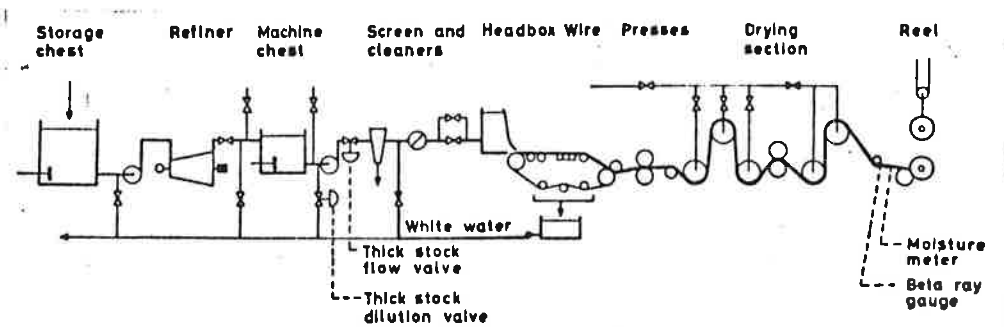
$$u(t) = -B^{-1}(q^{-1})\, G(q^{-1})\, F^{-1}(q^{-1})\, y(t)$$

$$= -B^{-1}(q^{-1})\, G(q^{-1})\, \varepsilon(t)$$

$$y(t) = F(q^{-1})\, \varepsilon(t)$$

$$A^{-1}(q^{-1})\, C(q^{-1}) = F(q^{-1}) + q^{-k} A^{-1}(q^{-1})\, G(q^{-1})$$

BASIS WEIGHT CONTROL OF PAPERMACHINE



SECOND ORDER MODEL

TWO TIME DELAYS

SEVEN PARAMETERS

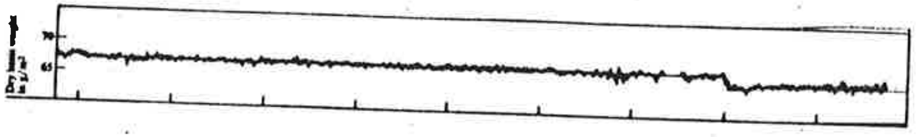
$$\Delta y(t) = \frac{4.61q - 4.05}{q^2 - 1.283q + 0.495}\,\Delta u(t-2) +$$

$$+\ 0.382\,\frac{q^2 - 1.438q + 0.550}{q^2 - 1.283q + 0.495}\, e(t)$$

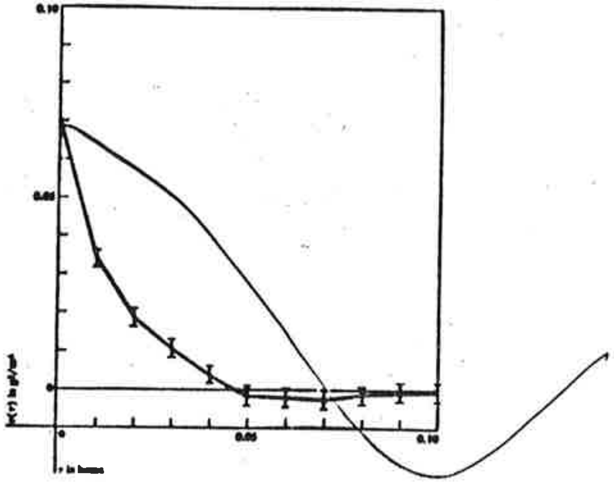
REF K. J. Å. INTRODUCTION TO STOCHASTIC CONTROL THEORY

ASSESSMENT OF NEED FOR RETUNING OF MINIMUM VARIANCE CONTROL

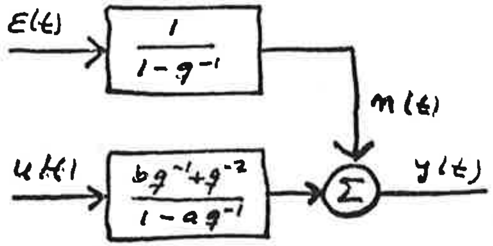
LOG CONTROLLED OUTPUT DURING NORMAL OPERATION



CALCULATE COVARIANCE FUNCTION OF OUTPUT
(RESID Y)



USE KNOWLEDGE OF PROCESS DEAD TIME AND SAMPLING PERIOD TO TEST IF COVARIANCE FUNCTION SATISFIES MINIMUM VARIANCE CONDITIONS



$$y(t) = \frac{b+q^{-1}}{1-aq^{-1}}\, u(t-1) + \frac{1}{1-q^{-1}}\, \varepsilon(t)$$

$$E\,\varepsilon^2(t) = 1 \,, \quad b \neq 0$$

CASE		Ey^2
STABLE MIN. PHASE	$\lvert a\rvert<1, \lvert b\rvert>1$	1
UNSTABLE MIN PHASE	$\lvert a\rvert>1, \lvert b\rvert>1$	a^2
STABLE NON MIN PHASE	$\lvert a\rvert<1 \quad \lvert b\rvert<1$	$\frac{2}{1+b}$
UNSTABLE NON MIN PHASE	$\lvert a\rvert>1 \quad \lvert b\rvert<1$	d

$$d = a^2\left[1 + \frac{(f_1-b)^2}{1-b^2}\right]$$

$$f_1 = \frac{a + a^2 - 1 + ab(a+1) + b^2}{a[1+b+ab+ab^2]}$$

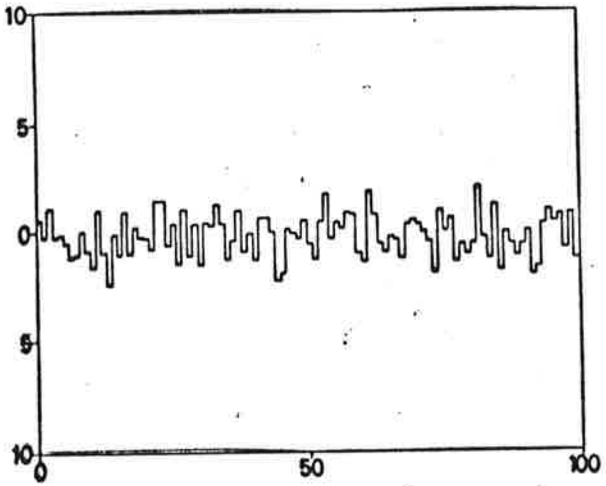
STABLE NON MIN PHASE SYSTEM

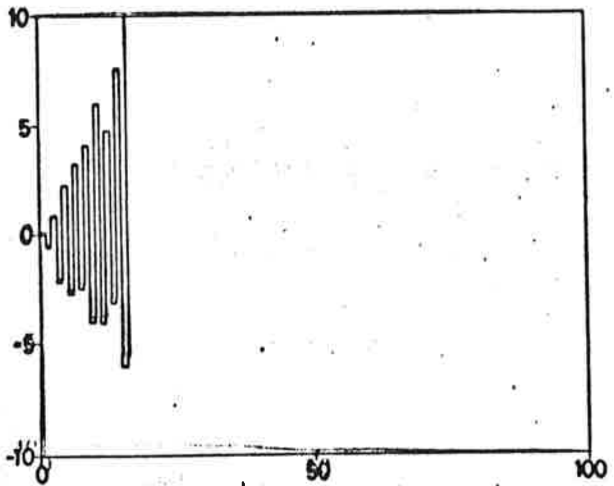
MIN VAR STRATEGY

$$A(q^{-1}) = 1 - 1.7q^{-1} + 0.7q^{-2}$$

$$B(q^{-1}) = 0.9 + q^{-1}$$

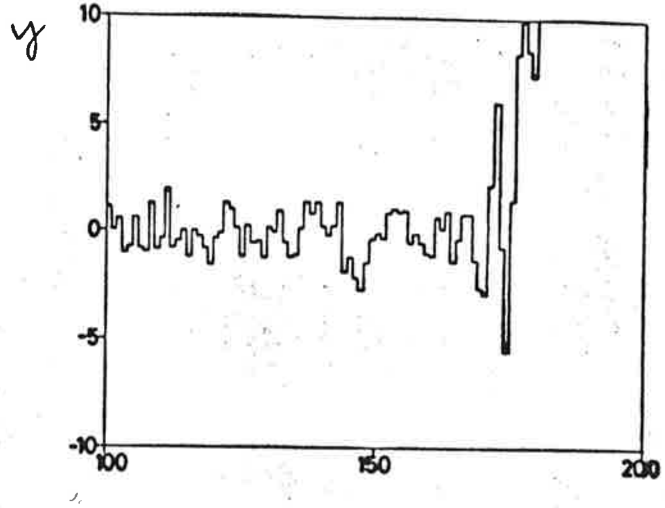
$$C(q^{-1}) = 1 - 0.7q^{-1}$$

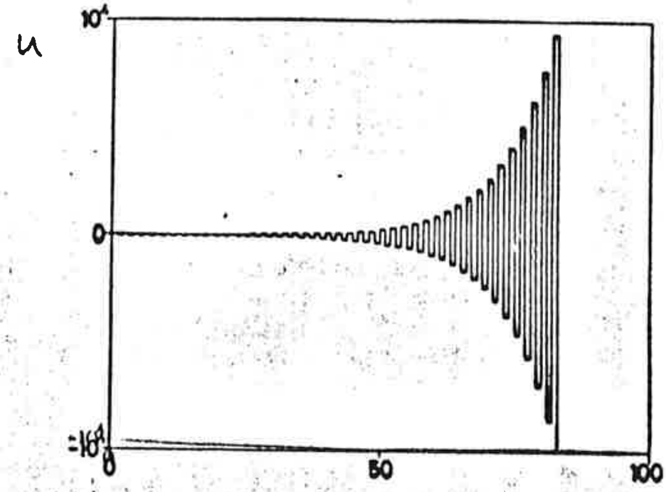


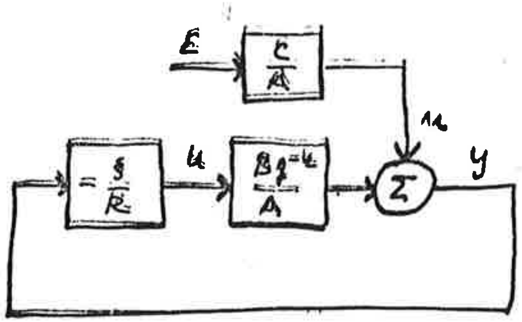


STABLE NON MIN PHASE SYSTEM

MIN VAR STRATEGY







$$y = \frac{AR}{AR+BSq^{-k}} \cdot \frac{C}{A}\varepsilon = \left[1 - \frac{q^{-k}BS}{AR+q^{-k}BS}\right]\frac{C}{A}\varepsilon$$

$$= F\varepsilon + q^{-k}\frac{G}{A}\varepsilon - q^{-k}\frac{BS}{AR+q^{-k}BS}\cdot\frac{C}{A}\varepsilon$$

$$= F\varepsilon + \left[1 - \frac{BSC}{G(AR+q^{-k}BS)}\right]q^{-k}\frac{G}{A}\varepsilon$$

$$= F\varepsilon(t) + [1 - BX]\frac{G}{A}\varepsilon(t-k)$$

$$E\,y^2 = E\,(F\varepsilon)^2 + J$$

$$J = \frac{1}{2\pi i}\oint \left\{[1-BX]\frac{G}{A}\right\}(z^{-1})\left\{[1-BX]\frac{G}{A}\right\}(z)\,\frac{dz}{z}$$

$$= \frac{1}{2\pi i}\oint [1-B(z^{-1})X(z^{-1})]\,Q(z)\,[1-B(z)X(z)]\,\frac{dz}{z}$$

J MAY HAVE SEVERAL LOCAL MIN

EXAMPLE

$$J[f] = \frac{1}{2\pi i}\oint [1-(z-a)f(z)][1-(z^{-1}-a)f(z)]\,\frac{dz}{z}$$

J min for

$$f(z) = \frac{1}{z-a} \qquad J[f] = 0$$

But if $|a| > 1$ local minimum also for

$$f(z) = \frac{1}{a^2}\cdot\frac{1}{z-1/a} \qquad J[f] = 1 - \frac{1}{a^2}$$

POLYNOMIAL B UNSTABLE

$$B = B^+ B^-$$

B^-: ALL ZEROS INSIDE UNIT DISC

B^+: ALL ZEROS OUTSIDE UNIT DISC

$$y = \frac{B}{A}u + \frac{C}{A}e = \frac{B^-}{\tilde{B}^-}\underbrace{\left[\frac{B^+\tilde{B}^-}{A}u + \frac{C\tilde{B}^-}{AB^-}e\right]}_{w}$$

$$Ey^2 = Ew^2$$

$$w = \frac{B^+\tilde{B}^-}{A}u + \frac{F}{B^-}e + \frac{G}{A}e$$

$$\deg F < \deg B^-$$

$$\frac{F}{B^-}e = f_0 e(t+1) + f_1 e(t+2) + \ldots$$

$$w = \frac{F}{B^-}e + \frac{1}{A}\left[B^+\tilde{B}^- u + Ge\right]$$

$$w = \frac{F}{B^-}e \Rightarrow y = \frac{F}{\tilde{B}^-}e$$

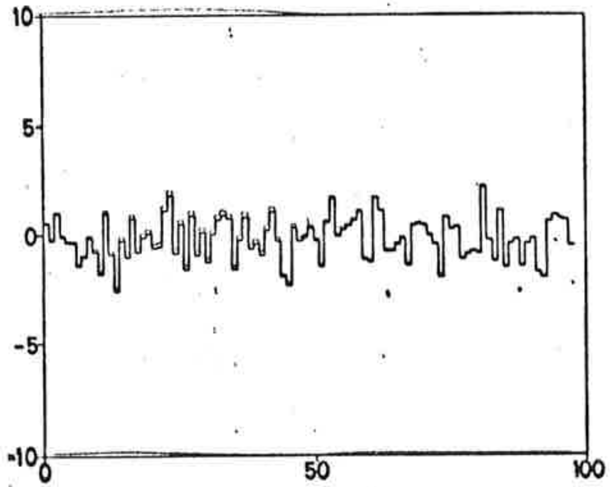
$$u = -\frac{G}{B^+\tilde{B}^-}e = -\frac{G}{B^+F}y$$

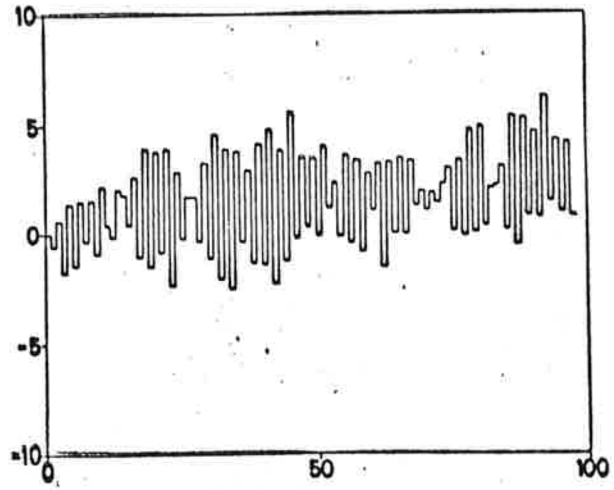
$$C\tilde{B}^- = AF + B^-G \qquad \deg F < \deg B^-$$

STABLE NON MIN PHASE SYSTEM

CONSTRAINED MIN VAR STRATEGY

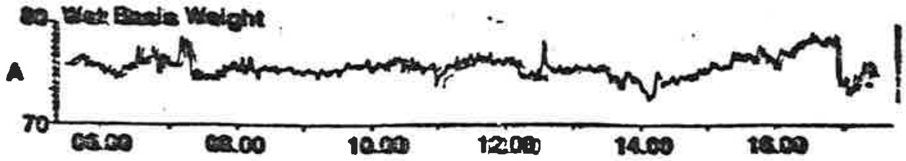
$\Sigma y^2(t) = 98.1 \qquad \Sigma u^2(t) = 690$





ASSESSMENT OF BENEFITS OF CONTROL

DATALOGGING:

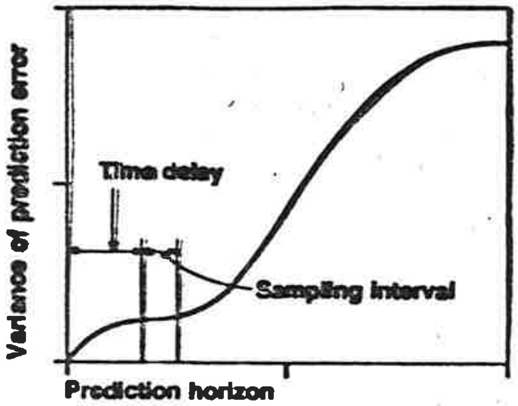


PROCESS IDENTIFICATION: PROCESS MODEL

PREDICTION

ERROR ANALYSIS

HEDGE



STOCHASTIC CONTROL THEORY AND APPLICATIONS

1. INTRODUCTION
2. LINEAR THEORY
   - INPUT - OUTPUT MODELS
   - STATE MODELS
3. MODELING & IDENTIFICATION
4. AN APPLICATION
5. NONLINEAR THEORY
6. SELF-TUNING CONTROL
7. AN APPLICATION
8. CONCLUSIONS

THE LQG PROBLEM
INTERNAL DESCRIPTIONS

1. MATHEMATICAL MODELS

2. OPTIMAL PREDICTION

   Kalman Filtering
   Innovations
   Duality

3. OPTIMAL CONTROL
   Loss Functions
   The Separation Principle

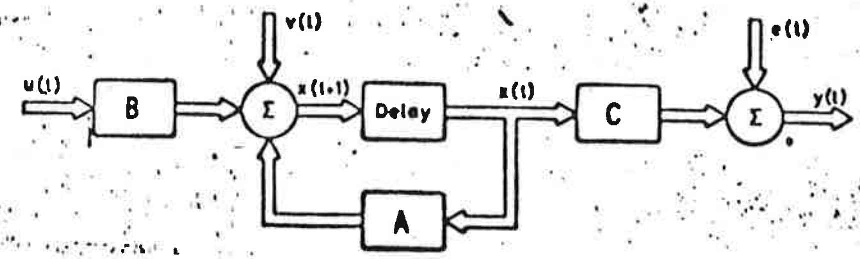
4. RELATIONS TO MIN. VAR. CONTROL
   Internal vs external descriptions
   Matrices vs rational functions
   Riccati Equations vs spectral factorizations

5 APPLICATIONS

2. LINEAR QUADRATIC CONTROL

$$x(t+1) = A\,x(t) + B\,u(t) + v(t)$$
$$y(t) = C\,x(t) + e(t)$$

$$J = \min E\left[x^T(t_1)\,Q_0\,x(t_1) + \sum x^T(t)\,Q_1\,x(t) + u^T(t)\,Q_2\,u(t)\right]$$



TWO SETS OF HYPOTHESES

- {e(t)} {v(t)} NORMAL RANDOM PROCESSES
  u(t) AN ARBITRARY FUNCTION OF y(t − 1), y(t − 2), ...
- {e(t)} {v(t)} SECOND ORDER RANDOM PROCESSES
  u(t) A LINEAR FUNCTION OF y(t − 1), y(t − 2), ...

KALMAN FILTERING

$$x(t+1) = Ax(t) + v(t) \qquad (+F(y_t))$$
$$y(t) = Cx(t) + e(t)$$

$\begin{Bmatrix} v(t) \\ e(t) \end{Bmatrix}$ independent gaussian $\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}R_1 & 0\\ 0 & R_2\end{bmatrix}\right)$

$x(t_0)$ gaussian (m, R_0)

Th. (Kalman 1960)

fins
normality
independent
Computations
sufficient. stat.
Practical base.
Asymptotic properties

The conditional distribution of $x(t+1)$ given Y_t is gaussian $[\hat{x}(t+1), P(t+1)]$ where

$$\hat{x}(t+1) = A\hat{x}(t) + K(t)[y(t) - C\hat{x}(t)] \qquad (+F(y_t))$$

$$K(t) = AP(t)C^T[CP(t)C^T + R_2]^{-1}$$

$$P(t+1) = AP(t)A^T + R_1 - AP(t)C^T[CP(t)C^T + R_2]^{-1}CP(t)A^T$$

$$= [A - K(t)C]P(t)A^T + R_1$$

$$= [A - K(t)C]P(t)[A - K(t)C]^T + R_1 + K(t)R_2K^T(t)$$

$$P(t_0) = R_0, \quad \hat{x}(t_0) = m$$

Proof.

$\begin{bmatrix} x \\ y \end{bmatrix}$ gaussian $\begin{bmatrix} m_x \\ m_y \end{bmatrix} \begin{bmatrix} R_x & R_{xy} \\ R_{yx} & R_y \end{bmatrix}$

$\Rightarrow E[x \mid y] = m_x + R_{xy}R_y^{-1}(y - m_y)$

INNOVATIONS REPRESENTATION OF $\{y(t)\}$

$$\begin{cases} x(t+1) = Ax(t) + v(t) \\ y(t) = Cx(t) + e(t) \end{cases}$$

Introduce

$$\varepsilon(t) = y(t) - C\hat{x}(t) = y(t) - \hat{y}(t)$$

$\varepsilon(t)$ is independent of $\varepsilon(s)$ for $t \neq s$

$$\hat{x}(t+1) = A\hat{x}(t) + K[y(t) - C\hat{x}(t)]$$

$$\begin{cases} \hat{x}(t+1) = A\hat{x}(t) + K\varepsilon(t) \\ y(t) = C\hat{x}(t) + \varepsilon(t) \end{cases}$$

Alternative representation for $\{y(t)\}$

$$E\,\varepsilon^2(t) = R_2 + CP(t)C^T$$

DUALITY

$$x(t+1) = Ax(t) + v(t) \quad \rightarrow R_1$$
$$y(t) = Cx(t) + e(t) \quad R_2$$

Predict $a^T x(t_1+1)$ linearly from $y(t_1), \dots, y(t_0)$

$$a^T x(t_1+1) = -\sum_{t_0}^{t_1} u^T(t)\, y(t) + b^T m$$

Introduce

$$z(t) = A^T z(t+1) + C^T u(t+1), \quad z(t_1) = a$$

$$a^T x(t_1+1) = z^T(t_1)\, x(t_1+1) = z^T(t_0-1)\, x(t_0)$$
$$+ \sum_{t_0}^{t_1} [z^T(t)\, x(t+1) - z^T(t-1)\, x(t)]$$
$$= z^T(t_0-1)\, x(t_0) + \sum_{t_0}^{t_1} z^T(t)\, v(t) - u^T(t)\, C x(t)$$

$$a^T x(t_1+1) - a^T \hat{x}(t_1+1) = z^T(t_0-1)\, x(t_0) - b^T m$$
$$+ \sum_{t_0}^{t_1} z^T(t)\, v(t) + u^T(t) \underbrace{[y(t) - Cx(t)]}_{e(t)}$$

$$E\{a^T[x(t_1+1) - \hat{x}(t_1+1)]\}^2 = z^T(t_0-1)\, R_0\, z(t_0-1)$$
$$+ \sum_{t_0}^{t_1} z^T(t)\, R_1\, z(t) + u^T(t)\, R_2\, u(t)$$

$$x(t+1) = Ax(t) + Bu(t) + v(t) \qquad x(t_0) \in N(m, R_0)$$
$$y(t) = Cx(t) + e(t)$$
$$J = E\left\{x^T(N)\, Q_0\, x(N) + \sum_{t_0}^{N-1} x^T(s)\, Q_1\, x(s) + u^T(s)\, Q_2\, u(s)\right\}$$

$u(t) =$ fcn of $y(t-1), \dots$ [measurable w.r.t. $\mathcal{Y}_{t-1}$

THEOREM (SEPARATION THEOREM)

Let the difference equation

$$S(t) = A^T S(t+1) A + Q_1 - A^T S(t+1) B [Q_2 + B^T S(t+1) B]^{-1} B^T S(t+1) A$$
$$S(N) = Q_0$$

have a solution which is positive semidef for $t_0 \le t \le N$. Let $L(t)$ be defined by

$$L(t) = [Q_2 + B^T S(t+1) B]^{-1} B^T S(t+1) A$$

Then the control law

$$u(t) = -L(t)\, E[x(t) \mid \mathcal{Y}_{t-1}] = -L(t)\, \hat{x}(t \mid t-1)$$

minimizes J. The minimum is given by

$$\min J = m^T S(t_0)\, m + \operatorname{tr} S(t_0) R_0 + \sum_{t=t_0}^{N-1} \operatorname{tr} S(t+1) R_1$$
$$+ \sum_{t=t_0}^{N-1} \operatorname{tr} P(t)\, L^T(t)\, B^T S(t+1)\, A$$

LEMMA

Assume that the difference equation

$$S(t) = A^T S(t+1) A + Q_1 - A^T S(t+1) B [Q_2 + B^T S(t+1) B]^{-1} B^T S(t+1) A$$

$$S(N) = Q_0$$

has a solution which is pos. semi def. for $t_0 \le t \le N$. Then

$$J = x^T(N) Q_0 x(N) + \sum_{t=t_0}^{N-1} x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t)$$

$$u(t) = -L\,E[x(t) \mid y_{t-1}]$$

$$= x^T(t_0) S(t_0) x(t_0) + \sum_{t=t_0}^{N-1} [u(t) + Lx(t)]^T [Q_2 + B^T S(t+1) B] [u(t) + L x(t)]$$

$$+ \sum_{t=t_0}^{N-1} \{ v^T(t) S(t+1) [Ax(t) + Bu(t)] +$$

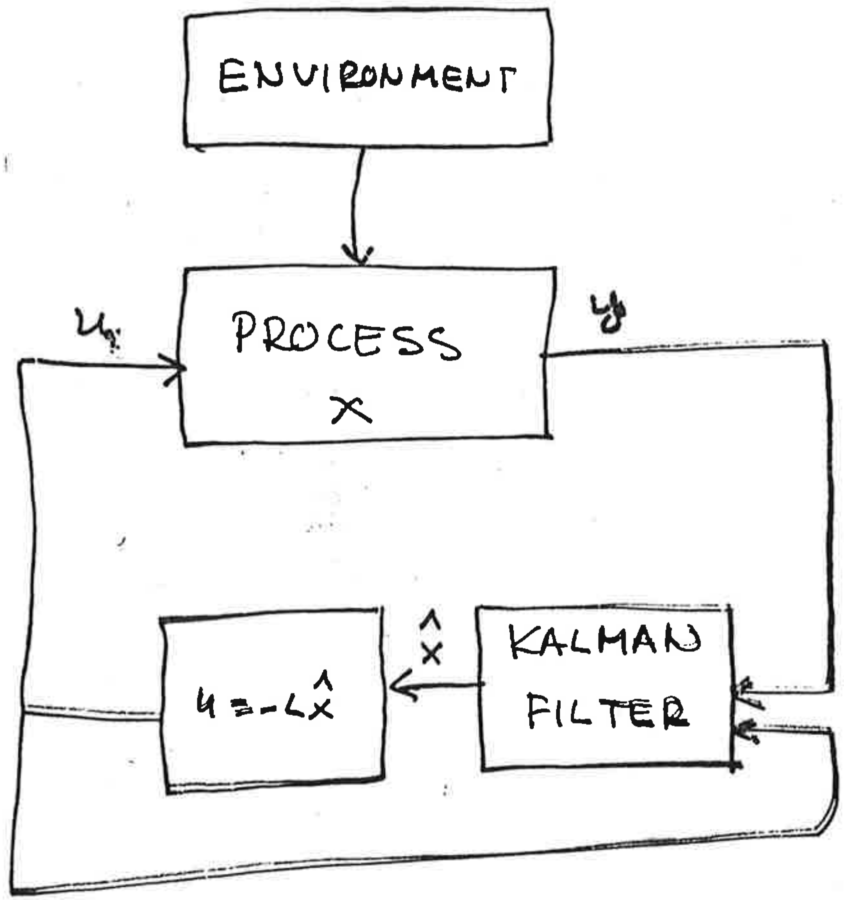
$$[Ax(t) + Bu(t)]^T S(t+1) v(t) + v^T(t) S(t+1) v(t)$$

Proof:

$$x^T(N) Q_0 x(N) = x^T(N) S(N) x(N) = x^T(t_0) S(t_0) x(t_0)$$

$$+ \sum_{t=t_0}^{N-1} [x^T(t+1) S(t+1) x(t+1) - x^T(t) S(t) x(t)]$$

STRUCTURE OF FEEDBACK GIVEN BY SEPARATION THEOREM



$$\begin{cases} \hat{x}(t+1) = A\hat{x}(t) + Bu(t) + K[y(t) - C\hat{x}(t)] \\ u(t) = -L\hat{x}(t) \end{cases}$$

SEPARATION THEOREM   CERTAINTY EQUIVALENCE PRINCIPLE

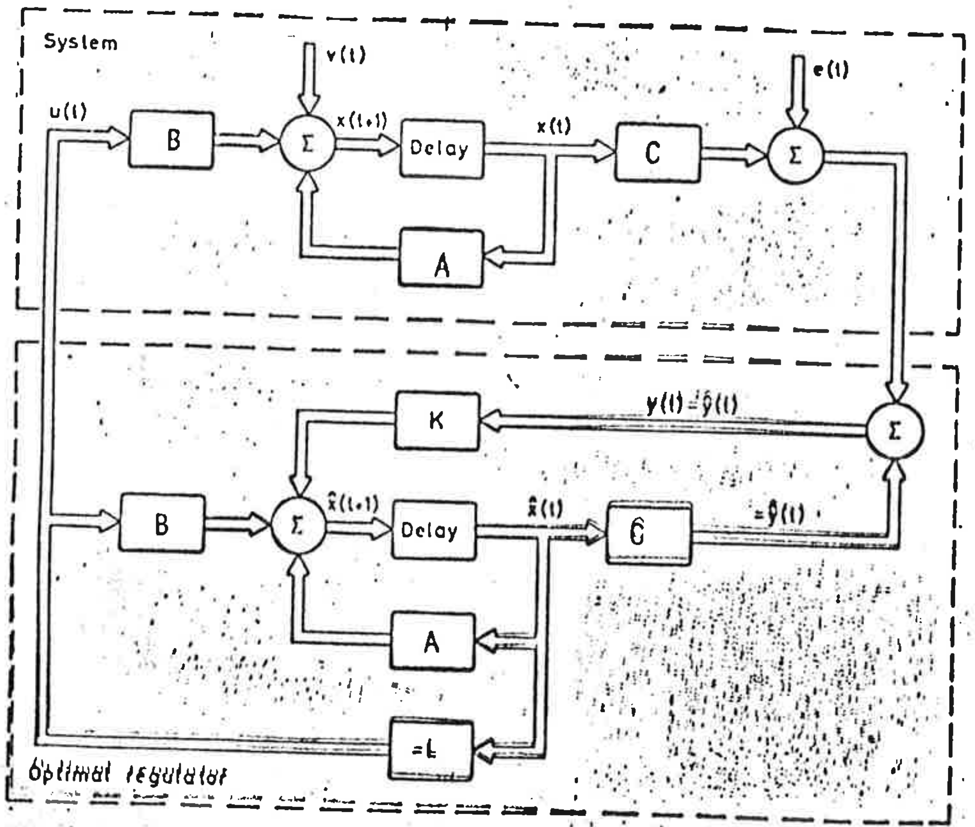
$$x(t+1) = A\,x(t) + B\,u(t) + v(t)$$

$$y(t) = C\,x(t) + e(t)$$

$$J = \min E\left[x^T(t_1)\, Q_0\, x(t_1) + \sum_{t=t_0}^{t_1-1} \left(x^T(t)\, Q_1\, x(t) + u^T(t)\, Q_2\, u(t) \right) \right]$$

$$u(t) = -L\,\hat{x}(t)$$

$$\hat{x}(t+1) = A\,\hat{x}(t) + B\,u(t) + K\left[y(t) - C\,\hat{x}(t)\right]$$



3. A SPECIAL CASE

$$x(t+1) = \begin{bmatrix} -a_1 & 1 \ldots 0 \\ -a_2 & 0 \ldots 0 \\ \vdots & \\ -a_{n-1} & 0 \ldots 1 \\ -a_n & 0 \ldots 0 \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} u(t) + \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_{n-1} \\ k_n \end{bmatrix} e(t)$$

$$y(t) = x_1(t) + e(t)$$

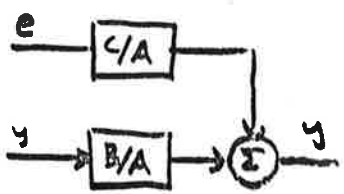
$$J = \min\ E\left[\frac{1}{N} \sum_{t=1}^{N} y^2(t)\right] \longrightarrow \min\ E\, y^2(t)$$

ONE STEP OPTIMIZATION THE SAME AS N STEP OPTIMIZATION

$$y(t) + a_1 y(t-1) + \ldots + a_n y(t-n) = b_1 u(t-1) + \ldots + b_n u(t-n) + \lambda\,[e(t) + c_1 e(t-1) + \ldots + c_n e(t-$$

$$A(q)\,y(t) = B(q)\,u(t-k) + \lambda\, C(q)\, e(t)$$

$$q\,y(t) = y(t+1)$$



EXAMPLE

$$y(t) + a y(t-1) = b u(t-1) + \lambda\,[e(t) + c e(t-1)]$$

$$u(t) = \frac{a - c}{b}\, y(t)$$

EXAMPLE

$$x(t+1) = \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t)$$

$$y(t+1) + a\,y(t) = b_1 u(t) + b_2 u(t-1)$$

$$J = \sum_{1}^{N} y^2(t) \quad \min \qquad N \to \infty$$

$$\begin{cases} S(t) = A^T S(t+1) A^T + Q_1 - A^T S(t+1) B [Q_2 + B^T S(t+1) B]^{-1} B^T S(t+1) A \\ S(N) = Q_0 \end{cases}$$

Solution 1

$$S(-\infty) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad L = \frac{1}{b_1}[-a, 1]$$

Solution 2

$$S = \begin{bmatrix} s_1 & s_2 \\ s_2 & s_3 \end{bmatrix}$$

$$s_1 = 1 + \frac{a^2(b_2^2 - b_1^2)}{(ab_1 - b_2)^2}$$

$$s_2 = -\frac{a(b_2^2 - b_1^2)}{(ab_1 - b_2)^2}$$

$$s_3 = \frac{b_2^2 - b_1^2}{(ab_1 - b_2)^2}$$

$$L = \frac{ab_2 - b_1}{b_2(ab_1 - b_2)}[-a, 1]$$

LQG EXTERNAL (INPUT-OUTPUT MODELS

GIVEN THE MODEL

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})e(t)$$

FIND CONTROL WHICH MINIMIZ

$$J = \lim_{N \to \infty} \frac{1}{N} \sum_{1}^{N} [y^2(t) + \rho u^2(t)]$$

1. SOLVE FACTORIZATION P

$$P\tilde{P} = \rho A\tilde{A} + B\tilde{B}$$

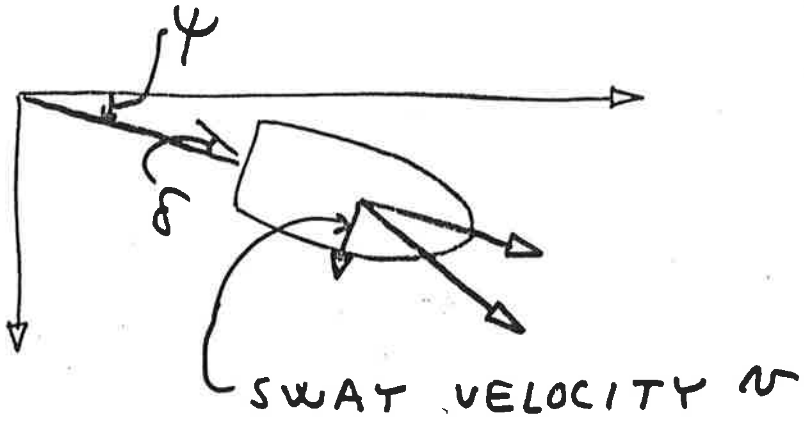
2. SOLVE DIOPHANTINE EQUATION

$$AR + BS = CP$$

3. OPTIMAL CONTROL LAW IS

$$R\,u(t) = C\,y_{ref}(t) - S\,y(t)$$

EXAMPLE SHIP STEERING



$$\frac{d}{dt}\begin{bmatrix} v \\ r \\ \psi \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} v \\ r \\ \psi \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{21} \\ 0 \end{bmatrix}\delta$$

$$\frac{\Delta R}{R} = k[\bar{\psi}^2 + \lambda\bar{\delta}^2]$$

TANKER $\qquad k = 0.014, \lambda = 0.1$

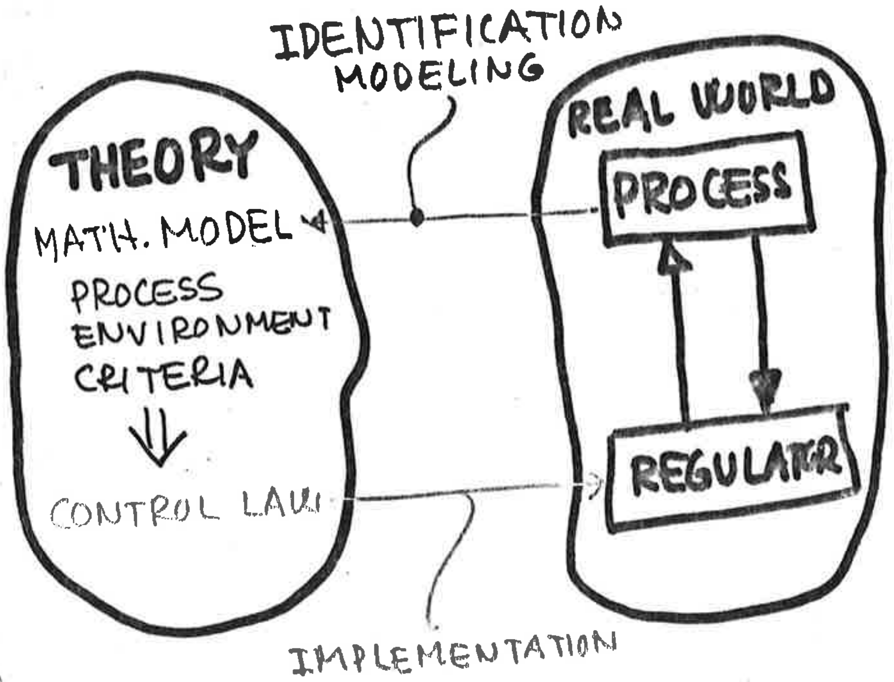
$$J = \lim_{T\to\infty} \frac{1}{T}\int_0^T [\psi^2(t) + \lambda\,\delta^2(t)]\,dt$$

STOCHASTIC CONTROL THEORY AND APPLICATIONS

1. INTRODUCTION
2. LINEAR THEORY
   - INPUT - OUTPUT MODELS
   - STATE MODELS
3. MODELING & IDENTIFICATION
4. AN APPLICATION
5. NONLINEAR THEORY
6. SELF-TUNING CONTROL
7. AN APPLICATION
8. CONCLUSIONS

MODELING AND IDENTIFICATION

1. INTRODUCTION
2. CRITERIA
3. DYNAMICAL SYSTEMS
4. MODEL STRUCTURES
5. ESTIMATION THEORY
6. INTERACTIVE COMPUTING
7. EXAMPLE
8. CONCLUSIONS



CLASSICAL:

TRANSFER FUNCTION
IMPULSE & FREQUENCY RESPONSE

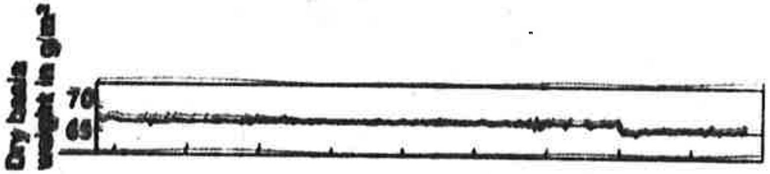
"MODERN":

PARAMETRIC STATE SPACE MODELS
:
LEAST SQUARES
MAXIMUM LIKELIHOOD
:

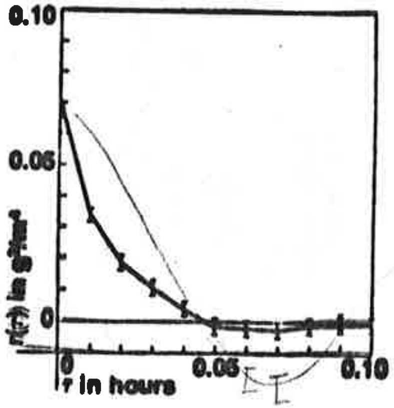
TRADE EXPERIMENTAL SIMPLICITY FOR COMPUTATIONS!

ASSESSMENT OF NEED FOR RETUNING OF MINIMUM VARIANCE CONTROL

LOG CONTROLLED OUTPUT DURING NORMAL OPERATION



CALCULATE COVARIANCE FUNCTION OF OUTPUT (COV Y)



USE KNOWLEDGE OF PROCESS DEAD TIME AND SAMPLING PERIOD TO TEST IF COVARIANCE FUNCTION SATISFIES MINIMUM VARIANCE CONDITIONS

MOTIVATION

PROCESS MODELLING

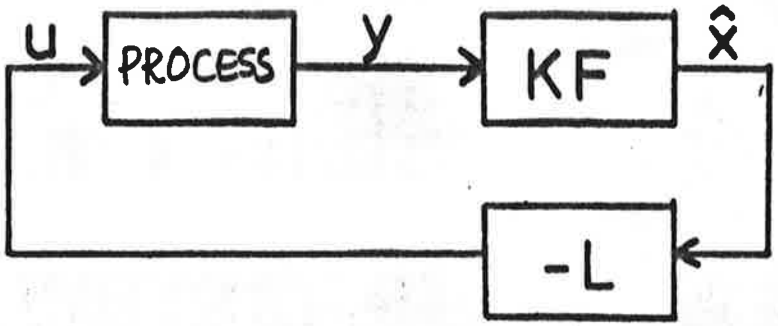
DESIGN OF CONTROL LAWS

Ex: Given the system

$$x(t+1) = Ax(t) + Bu(t) + V(t)$$
$$y(t) = Cx(t) + e(t)$$

Find control which minimizes

$$E \sum_{t=1}^{N} x^T(t)Q_1 x(t) + u^T(t)Q_2 u(t) \}$$



THEORIA MOTVS CORPORVM COELESTIVM

IN

SECTIONIBVS CONICIS SOLEM AMBIENTIVM

AVCTORE

CAROLO FRIDERICO GAVSS.

GÖTTINGEN 1809

The probability of the errors

$$\Omega = h^{\mu}\pi^{-\frac{1}{2}\mu}e^{-hh(vv+v'v'+v''v''+\dots)}$$

must become a minimum.

"Therefore, that will be the most probable system of values of the unknown quantities p, q, r, s, etc., in which the sum of the squares of the differences between the observed and computed values of the functions V, V', V'', etc. is a minimum, ..."

PRINCIPLE OF LEAST SQUARES

"In conclusion, the principle that the sum of the squares of the differences between the observed and computed quantities must be a minimum may, in the following manner, be considered independently of the calculus of probabilities."

"Denoting the differences between observation and calculation by Δ, Δ', Δ'', etc., the first condition will be satisfied not only if $\Delta\Delta + \Delta'\Delta' + \Delta''\Delta'' +$ etc., is a minimum (which is our principle), but also if $\Delta^4 + \Delta'^4 + \Delta''^4 +$ etc., or $\Delta^6 + \Delta'^6 + \Delta''^6 +$ etc., or in general, if the sum of any of the powers with an even exponent becomes a minimum. But of all these principles ours is the most simple; by the others we should be led into the most complicated calculations."

THE MAXIMUM LIKELIHOOD PRINCIPLE — Fisher 1912

RULE

Let Y be a random variable with probability density $p(y, \theta)$. To estimate θ from an observation y choose θ such that

$$L(y, \hat{\theta}) \geq L(y, \theta) \quad \forall \theta$$

where L is the likelihood function defined by $L(y, \theta) = p(y, \theta)$.

INDEPENDENT SAMPLES

$$L(y_1, y_2, \ldots y_n, \theta) = p(y_1, \theta)\, p(y_2, \theta) \ldots p(y_n, \theta)$$

PROPERTIES

Consistency

Asymptotic normality

Efficiency

THE LIKELIHOOD FUNCTION

INTRODUCE

$$y_{t_k}^T = \left[y^T(t_0)\ y^T(t_1) \ldots y^T(t_k) \right]$$

USING BAYES RULE THE LIKELIHOOD FUNCTION BECOMES

$$L = p(y_{t_k}) = p(y(t_k) \mid y_{t_{k-1}})\, p(y_{t_{k-1}})$$

$$p(y(t_k) \mid y_{t_{k-1}})\, p(y(t_{k-1}) \mid y_{t_{k-2}}) \ldots p(y(t_1) \mid y(t_0))\, p(y(t_0))$$

FOR LINEAR GAUSSIAN PROBLEMS

$$p(y(t_k) \mid y_{t_{k-1}}) = N(\hat{y}(t_k \mid t_{k-1}),\ R(t_k))$$

$$= (1/2)(2\pi)^{-m/2} (\det R(t_k))^{-1/2} \exp -(1/2)\, \varepsilon^T(t_k) R^{-1}(t_k) \varepsilon(t_k)$$

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k)$$

THE LIKELIHOOD FUNCTION CAN THEN BE WRITTEN AS

$$-\log L = (1/2) \left[\Sigma \log \det R(t_k) + \Sigma \varepsilon^T(t_k) R^{-1}(t_k) \varepsilon(t_k) \right] + \text{const.}$$

NOTICE RELATIONS TO FILTERING THEORY !

INTERPRETATION FOR NON GAUSSIAN PROCESSES

PREDICTION ERROR INTERPRETATION

Notice that the ML-criterion gives a loss function N of the form

$$V(\theta) = \sum_{t=1}^{N} q(\varepsilon(t_k))$$

where

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k|t_{k-1})$$

is the prediction error.

Alternative: Postulate prediction model and error criterion

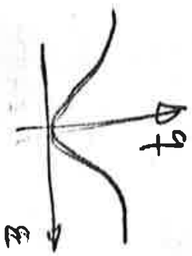
Compare Gauss

Consequences for modeling

Dynamical systems

NOTICE q QUADRATIC FOR GAUSSIAN DISTURBANCES

ROBUSTNESS HUBER



OTHER PREDICTION ERROR CRITERIA

ML: fcn of θ

$$V(\theta) = -\log L = \frac{1}{2}\sum_{k=1}^{N} \log\det R(t_k) + \frac{1}{2} n_y N \log 2\pi + \frac{1}{2}\sum_{k=1}^{N} \varepsilon^T(t_k) R^{-1}(t_k)\varepsilon(t_k)$$

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k|t_{k-1})$$

MORE GENERAL

$$V(\theta) = g(G(\theta))$$

$$G(\theta) = \sum_{k=1}^{N} F(\varepsilon(t_k), \theta, t_k)$$

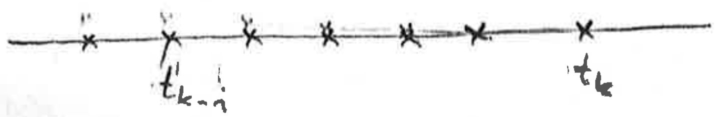
LONGER PREDICTION HORIZON

$$V(\theta) = g(G_1(\theta), G_2(\theta), \dots, G_s(\theta))$$

$$G_i(\theta) = \sum_{k=1}^{N} F_i(\varepsilon(t_k|t_{k-i}), \theta, t_k)$$

$$\varepsilon(t_k|t_{k-i}) = y(t_k) - \hat{y}(t_k|t_{k-i})$$

$$\vdots$$



ESTIMATING PARAMETERS OF DYNAMICAL SYSTEMS

Example

$$\dot{x} = Ax + Bu + \underset{R_1}{v}$$

$$y(t_k) = Cx(t_k) + \underset{R_2}{e(t_k)}$$

How to obtain the likelihood function

Computational aspects

The minimization problem

Properties of the ML-estimate

THE LIKELIHOOD FUNCTION

INTRODUCE

$$y_{t_k}^T = \left| y^T(t_0)\ y^T(t_1) \ldots y^T(t_k) \right|$$

USING BAYES RULE THE LIKELIHOOD FUNCTION BECOMES

$$L = p(y_{t_k}) = p(y(t_k)|y_{t_{k-1}})\ p(y_{t_{k-1}})$$

$$p(y(t_k)|y_{t_{k-1}})\ p(y(t_{k-1})|y_{t_{k-2}}) \ldots p(y(t_1)|y(t_0))\ p(y(t_0))$$

FOR LINEAR GAUSSIAN PROBLEMS

$$p(y(t_k)|y_{t_{k-1}}) = N(\hat{y}(t_k|t_{k-1}),\ R(t_k))$$

$$= (1/2)(2\pi)^{-m/2}(\det R(t_k))^{-1/2} \exp -(1/2)\,\varepsilon^T(t_k)R^{-1}(t_k)\varepsilon(t_k)$$

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k)$$

THE LIKELIHOOD FUNCTION CAN THEN BE WRITTEN AS

$$-\log L = (1/2)\left|\Sigma \log\det R(t_k) + \Sigma\, \varepsilon^T(t_k)R^{-1}(t_k)\varepsilon(t_k)\right| + \text{const.}$$

NOTICE RELATIONS TO FILTERING THEORY !

INTERPRETATION FOR NON GAUSSIAN PROCESSES

EXAMPLE

$$\dot{x} = Ax + Bu + v$$

$$y(t_k) = Cx(t_k) + e(t_k)$$

THE KALMAN BUCY THEORY GIVES:

$$\hat{y}(t_k|t_{k-1}) = C\hat{x}(t_k|t_{k-1})$$

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k|t_{k-1})$$

$$R(t_k) = R_2 + C\,P(t_k|t_{k-1})C^T$$

$$\hat{x}(t_k|t_k) = \hat{x}(t_k|t_{k-1}) + K(t_k)\cdot\varepsilon(t_k)$$

$$K(t_k) = P(t_k|t_{k-1})C^T R^{-1}(t_k)$$

$$P(t_k|t_k) = P(t_k|t_{k-1}) - K(t_k)CP(t_k|t_{k-1})$$

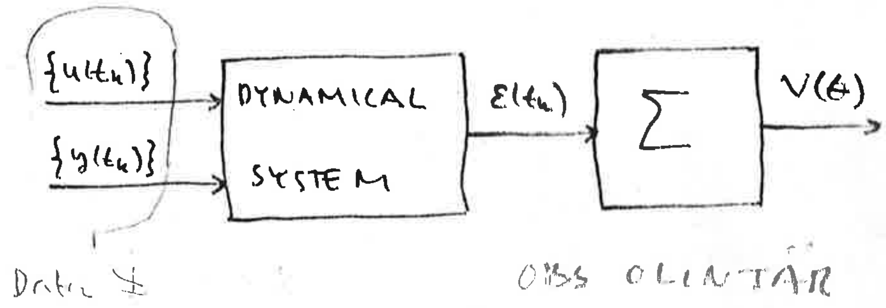
$$\frac{d\hat{x}(t|t_k)}{dt} = A\hat{x}(t|t_k) + Bu(t) \qquad t_k \leq t \leq t_{k+1}$$

$$\frac{dP(t|t_k)}{dt} = AP(t|t_n) + P(t|t_k)A^T + R_1 \qquad t_k \leq t \leq t_{k+1}$$

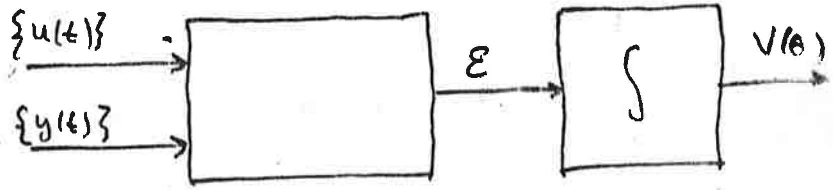
THE LIKELIHOOD FUNCTION

$$(-2\log L)_{t_k} = (-2\log L)_{t_{k-1}} + \varepsilon^T(t_k)R^{-1}(t_k)\varepsilon(t_k) + \log\det R(t_k)$$

NOTICE THE STRUCTURE OF THE LIKELIHOOD FUNCTION



CONTINUOUS TIME DATA



$$\frac{dz}{dt} = F(z, u(t), y(t), t) = G(z, t)$$

$$\varepsilon = H(z, t)$$

$$V(\theta) = \frac{1}{2}\int_0^T K(\varepsilon, t, \theta)\,dt$$

COMPUTATIONAL ASPECTS

What must be done?

Minimization algorithms → FUNCTION EVALUATION
GRADIENT -"-
HESSIAN -"-

Simplifications

constant sampling rate

special model structures

EXAMPLE $t_{k+1} - t_k = 1$

$$x(t+1) = A\,x(t) + B\,u(t) + K\,\varepsilon(t)$$

$$y(t) = C\,x(t) + \varepsilon(t)$$

$$-2 \log L = \sum_{1}^{N} \varepsilon^T(t)\, R^{-1}\, \varepsilon(t) + N \log \det R + C$$

MINIMIZE W.R.T. R!

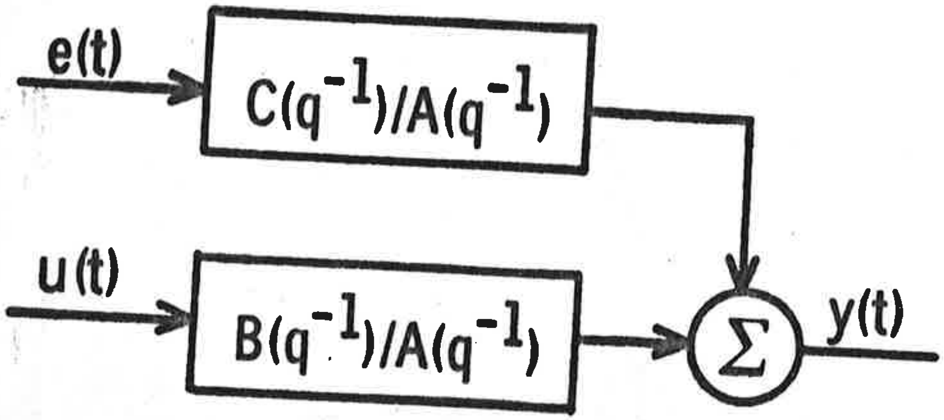
$$-2 \log L = N \log \det \frac{1}{N} \sum_{1}^{N} \varepsilon^T(t)\, \varepsilon^T(t) + r N + \text{const.}$$

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) + K\left[y(t) - Cx(t)\right] \\ \varepsilon(t) = y(t) - Cx(t) \end{cases}$$

$$\hat{R} = \frac{1}{N} \sum_{1}^{N} \varepsilon(t)\, \varepsilon^T(t)$$

EXAMPLE (ARMAX MODEL)

$$y(t) + a_1 y(t-1) + \ldots + a_n y(t-n) =$$
$$= \qquad b_1 u(t-1) + \ldots + b_n u(t-n) +$$
$$+ \lambda\left(e(t) + c_1 e(t-1) + \ldots + c_n e(t-n)\right)$$



$$A(q^{-1})\, y(t) = B(q^{-1})\, u(t) + C(q^{-1})\, e(t)$$
$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n}$$
$$B(q^{-1}) = \qquad b_1 q^{-1} + \ldots + b_n q^{-n}$$
$$C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_n q^{-n}$$

MINIMIZATION

$$-\log L = \frac{1}{\lambda^2} V(\theta) + \frac{N}{2} \log \lambda + \text{const}$$

$$V(\theta) = \frac{1}{2} \sum_{t=1}^{N} \varepsilon^2(t)$$

$$C(q^{-1})\varepsilon(t) = A(q^{-1})y(t) - B(q^{-1})u(t)$$

$$\theta^{k+1} = \theta^k - \left[V_{\theta\theta}(\theta^k)\right]^{-1} V_\theta(\theta^k)$$

$$V_\theta = \sum_{t=1}^{N} \varepsilon(t)\, \varepsilon_\theta(t)$$

$$V_{\theta\theta} = \sum_{t=1}^{N} \varepsilon_\theta(t)\varepsilon_\theta(t) + \sum_{t=1}^{N} \varepsilon(t)\varepsilon_{\theta\theta}(t)$$

$$C(q^{-1}) \frac{\partial \varepsilon(t)}{\partial a_i} = y(t-i)$$

$$C(q^{-1}) \frac{\partial \varepsilon(t)}{\partial b_i} = -u(t-1)$$

$$C(q^{-1}) \frac{\partial \varepsilon(t)}{\partial c_i} = -\varepsilon(t-i)$$

THE ARMAX MODEL

CANONICAL FORM FOR LINEAR TIMEINVARIANT SYSTEM WHOSE DYNAMIS IS RATIONAL TRANSFER FUNCTION + TIME DELAY

DISTURBANCES ARE STATIONARY WITH RATIONAL SPECTRAL DENSITY

CAN BE EXTENDED TO MISO!

$$Ay = B_1 u_1 + B_2 u_2 + \dots + B_r u_r + Ce$$

USING ADJOINT VARIABLES TO CALCULATE GRADIENTS

$$\frac{dx}{dt} = f(x,\theta,t)$$

$$V(\theta) = \int_0^T g(x,s)\,ds$$

$$V_\theta(\theta) = \int_0^T g_x x_\theta \, ds = -\int_0^T p^T(s) f_\theta \, ds$$

$$\begin{cases} \dfrac{dp}{dt} = -\left(\dfrac{\partial f}{\partial x}\right)^T p + g_x^T \\ p(T) = 0 \end{cases}$$

PROOF:

$$\frac{dx_\theta}{dt} = f_x x_\theta + f_\theta$$

$$V_\theta = \int_0^T [g_x x_\theta + p^T \dot{x}_\theta - p^T f_x x_\theta - p^T f_\theta]\,ds$$

$$= p^T x_\theta \Big]_0^T + \int_0^T [g_x x_\theta - \dot{p}^T x_\theta - p^T f_x x_\theta - p^T f_\theta]\,ds$$

$$= p^T x_\theta \Big]_0^T - \int_0^T [g_x - p^T f_x - \dot{p}^T] x_\theta - \int_0^T p^T f_\theta \, ds$$

MODEL STRUCTURES

$$dx = Ax\,dt + Bu\,dt + dw \leftarrow R_1$$

$$dy = Cx\,dt + de \leftarrow R_2$$

$$y(t) + A_1 y(t-1) + \ldots + A_n y(t-n) =$$
$$= B_1 u(t) + \ldots + B_n u(t-n) + e(t) + C_1 e(t-1) + \ldots + C_n e(t-n)$$

$$y(t) = H(s)\,u(t) + G(s)\,e(t)$$

NONLINEAR MODELS

$$\frac{d\hat{x}(t|t_k)}{dt} = f(\hat{x}(t|t_k), u(t))$$

$$\hat{y}(t_k|t_{k-1}) = g(\hat{x}(t_k|t_{k-1})) +$$

$$\hat{x}(t_k|t_k) = h(\hat{x}(t_k|t_{k-1}), \varepsilon(t_k))$$

$$\varepsilon(t_k) = y(t_k) - \hat{y}(t_k|t_{k-1})$$

ESTIMATION THEORY

HOW WILL THE METHODS WORK
UNDER IDEAL CIRCUMSTANCES

HOW ARE THE RESULTS INFLUENCED
BY DIFFERENT CHOICES OF THE
PROBLEM ELEMENTS $\mathcal{D}$, $\mathcal{M}$, $\mathcal{C}$

CLASSICAL STATISTICS

- CONSISTENCY
- ASYMPTOTIC DISTRIBUTIONS
- EFFICIENCY

GENERAL COMMENT ON RESULTS

- LARGE SAMPLE PROPERTIES $N \to \infty$
- CHARACTER OF RESULTS

NOTIONS

$\mathcal{D}$ DATA GENERATED ~~FROM~~ M_0

$\mathcal{M}$ MODEL SET

$\mathcal{C}$ CRITERIA

INTRODUCE

$$W(\theta) = \lim_{N\to\infty} \frac{1}{N} V_N(\theta) = \left[-\lim_{N\to\infty} \frac{1}{N} \log L(\theta, y_N)\right]$$

SHOW <u>UNIFORM</u> CONVERGENCE
(ERGODIC THEOREMS OR MARTINGALE THEOREMS)

ANALYSE $W(\theta)$ FIND θ_0 WHICH
MINIMIZES $W(\theta)$

UNDER GENERAL BUT MESSY CONDITIONS

$$\hat{\theta}_N \to \theta_0$$

AVAILABLE COMMANDS IDPAC

1. INPUT & OUTPUT
   CONV, EDIT, MOVE, LIST
2. DISPLAY
   PLOT, BODE, PLMAG, FHEAD
3. DATA OPERATIONS
   INSI, CUT, CONC, PICK, SLIDE, STAT
   SCLOP, VECOP, TREND, ACOF, CCOF
4. FREQUENCY RESPONSE
   FROP, ASPEC, CSPEC, SPTRF
5. SIMULATION & MODEL ANALYSIS
   FILT, DSIM, DETER, RESID, RALIPA
6. IDENTIFICATION
   ML, STRUC, SQR, LS

⋮

9. MISC
   DELETE, TURN

AVAILABLE COMMANDS INTRAC

MACRO NAME [FARG1 FARG2 ...]
FORML FARG1 [FARG2 ...]
END
EXEC 'ON'/'OFF'
LABEL LNAM
GO TO LNAM
IF ARG1 RELOP ARG2 GOTO LAM
FOR COUNT = BEGIN TO FINISH [STEP INCR]
NEXT COUNT
SWITCH
LET VAR1 = ARG1 [OP ARG2]
READ VAR1 TYPE1 [VAR2 TYPE2]
WRITE [DEV] [STRING1/ARG1] ...
FREE
STOP

LS

$$A(q^{-1})y(t) = B_1(q^{-1})u_1(t-k_1) + \dots + e(t)$$

$$y(t) - \phi(t)\theta = e(t)$$

$$Y - \Phi\Theta = E \Rightarrow [\Phi \; Y]\begin{bmatrix} -\Theta \\ 1 \end{bmatrix} = E$$

$$Q[\Phi \; Y] = \begin{bmatrix} R \\ 0 \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n+1} \\ 0 & r_{22} & \dots & r_{2n+1} \\ \vdots & & & \\ 0 & & & r_{nn\,n+1} \\ 0 & \dots & 0 & r_{n+1n+1} \end{bmatrix} \begin{bmatrix} -\theta_1 \\ -\theta_2 \\ -\theta_n \\ 1 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_n \\ e_{n+1} \end{bmatrix}$$

SQR RMAT ← DATA STRF

LS [(SW)] SYST [(NAME)] ← STRF

SUBCOMMANDS

SAVE STDEV
SAVE COMAT

ML

IDENTIFICATION OF MODEL

$$A(q^{-1})y(t) = B_1(q^{-1})u_1(t) + \dots + B_m(q^{-1})u_m(t) + \lambda C(q^{-1})e(t)$$

ML [(SW)] SYST [(NAME)] ← DATA NO

SUBCOMMANDS

INVAL ABC/C SYST [(NAME)]
FIX A(1) [VAL] ...
SAVE [COMAT] [STDEV] [GRAD] [EVAL]
EXIT

PARAMETER ESTIMATES
-"- UNCERTAINTIES
LOSS

AKAIKE $AIC = -2 * \log ML + 2p$

CREATION OF NEW COMMAND

```
>MACRO LSID MODEL← DATA N
> STRUC STRF
    > NA MAX N
    > NU MAX 1
    > NB MAX N @
> SQR RMAT ← DATA STRF
> DELET MODEL
> LS MODEL ← STRF
> END
```

THE NEW COMMAND CAN BE USED AS FOLLOWS:

```
LSID ADAM ← DATA 2.
```

GENERATION OF COMMAND WITH A LOT OF GUIDANCE

```
>EDIT FILTR.
   NOT FOUND: FILTR
   INPUT
 MACRO FILTR.
 WRITE (TP) 'ENTER NAME OF FILTER'
 READ FILTN NAME
 WRITE (TP) 'WHAT TYPE (HP, BP, LP)
             IS 'FILTN' ?'
 READ FTYP NAME
 WRITE (TP) 'WHAT FILTER ORDER?'
 READ N INT
 IF FTYP EQ BP GOTO L1
 :
 LABEL L1
 WRITE (TP) 'ENTER LOW AND
    HIGH CUT-OFF FREQUENCIES (RAD/S)'
 READ LCF ~~LCF HCF~~ REAL HCF REAL
 FILT FILTN ← FTYP N LCF HCF
 END
 > EXIT
```

```
> FILTR
ENTER NAME OF FILTER
TEST
WHAT TYPE (HP, BP, LP) IS TEST?
BP
WHAT FILTER ORDER?
6
ENTER LOW AND HIGH CUT-OFF
FREQUENCIES (RAD/S)
5 30
>
```

PROGRAM DETAILS

52 SUBROUTINES 9200 STATEMENTS

64k UNIVAC 1108

25k USING SEGMENTATION

OPTIMIZATION ALGORITHMS

Quasi-Newton Fletcher

Brent

APPLICATIONS

SHIP STEERING DYNAMICS

INSULIN KINETICS

GLUCOSE KINETICS

ECONOMETRICS

CONCLUSIONS

- COMMANDS THAT ARE NATURAL FROM THEORY
- NEW COMMANDS CAN EASILY BE CREATED E.G. 2LS GLS ETC
- FLEXIBILITY
- SPECIAL PURPOSE TOOLS CAN EASILY BE IMPLEMENTED
- NOT ALL PROBLEMS ARE SUITABLE FOR INTERACTION
- COPY OF DIALOGUE
- DIRECT DOCUMENTATION
- "IDENTIFICATION LANGUAGE"

EXAMPLE ARMAX

SCENARIO : EXPLORATORY PHASE

$$A(q)y(t) = B(q)u(t) + \lambda C(q)e(t)$$

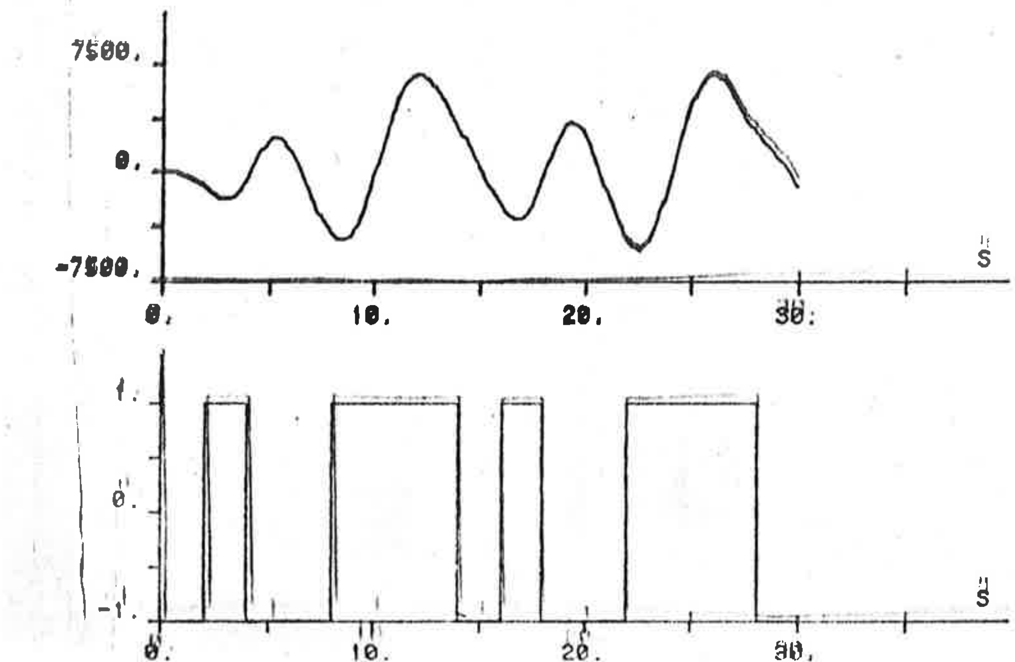
$$A(q) = (q^2 - 1.95q + 0.96)(q - 0.9)$$

$$\omega h = 0.1, \ \xi = 0.2$$

$$B(q) = q^2 + q + 1$$

$$C(q) = q^2 + 0.7q + 0.2$$

PLOT (30) YS/(HP) U



INTERACTIVE COMPUTING

IDPAC

MOVE WRK(1) ← U

MOVE WRK(2) ← YS

ML P1 ← WRK 1

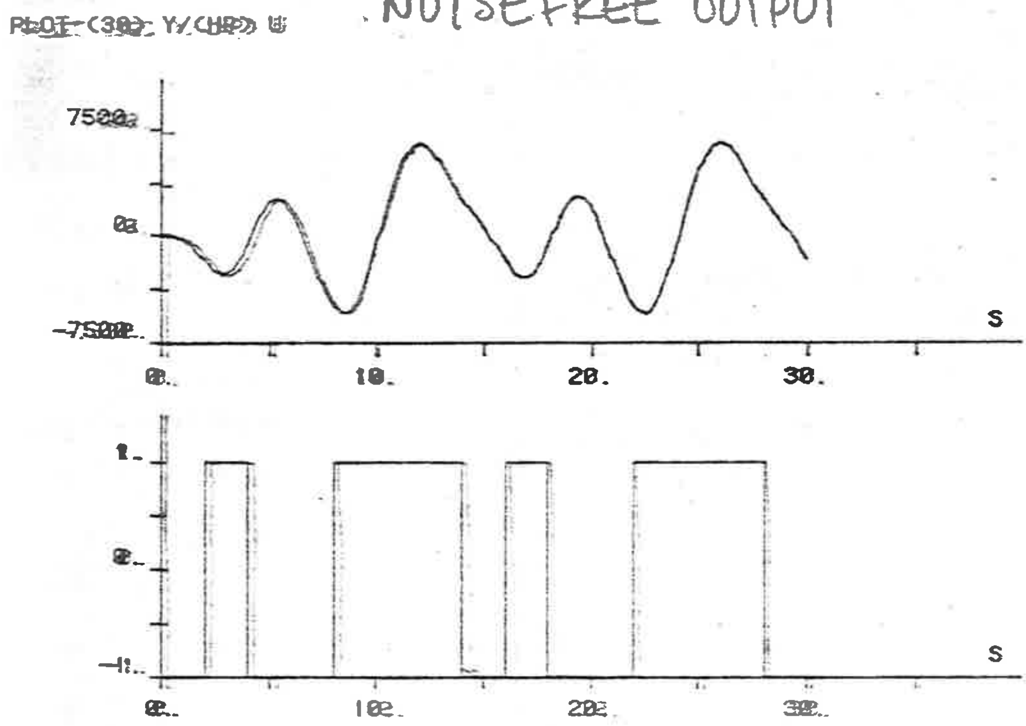
ML P2 ← WRK 2

ML P3 ← WRK 3

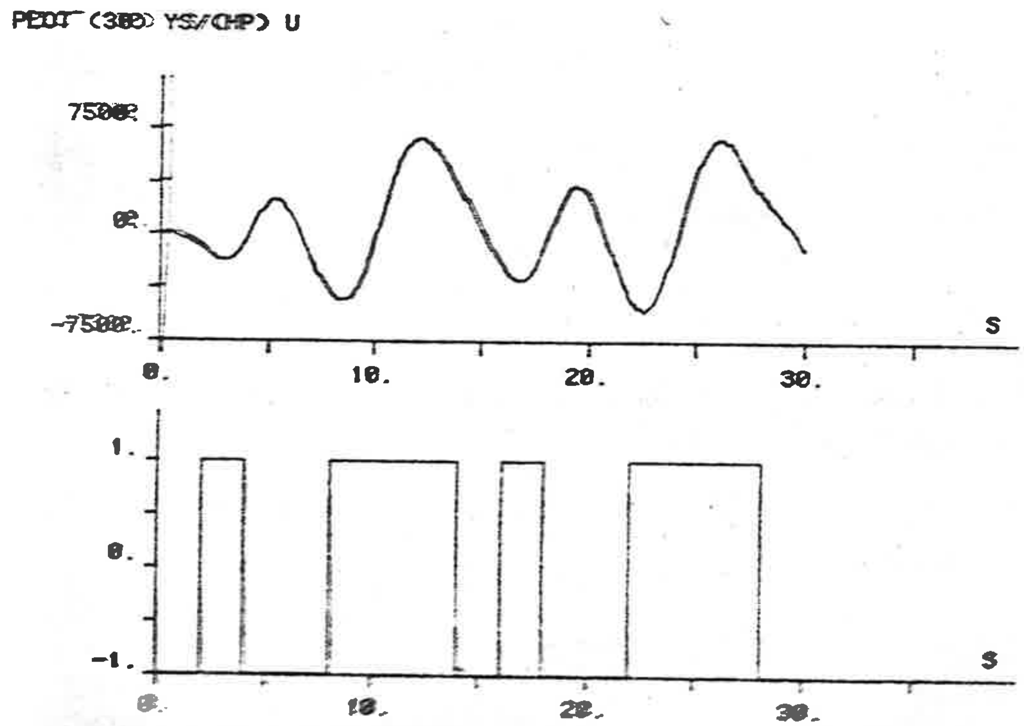
ML P4 ← WRK 4

n	V	AIC	A	B	C
1	2.4E6	3768	-0.995	62.1	1.00
2	172.8	1597	-1.979 0.985	4.90 4.37	1.66 1.79
3	139.3	847	-2.851 2.717 -0.865	1.06 0.81 1.05	0.72 0.20 0.03
4	138.0	850			

NOISEFREE OUTPUT

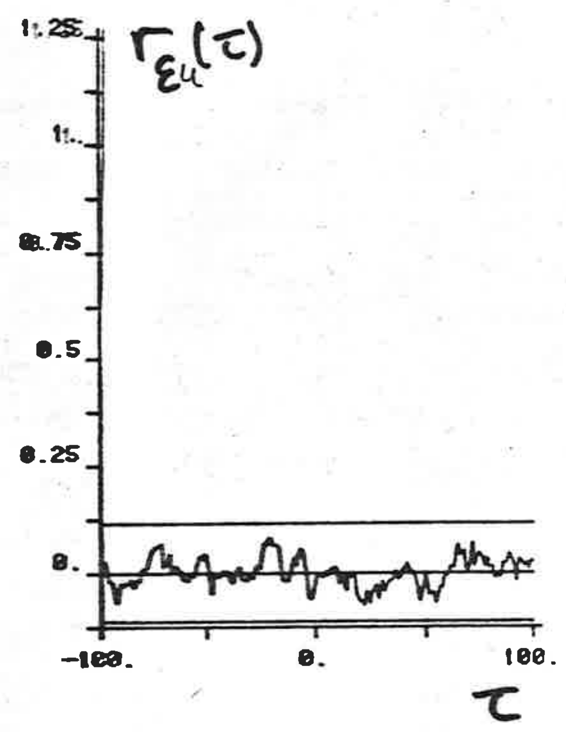


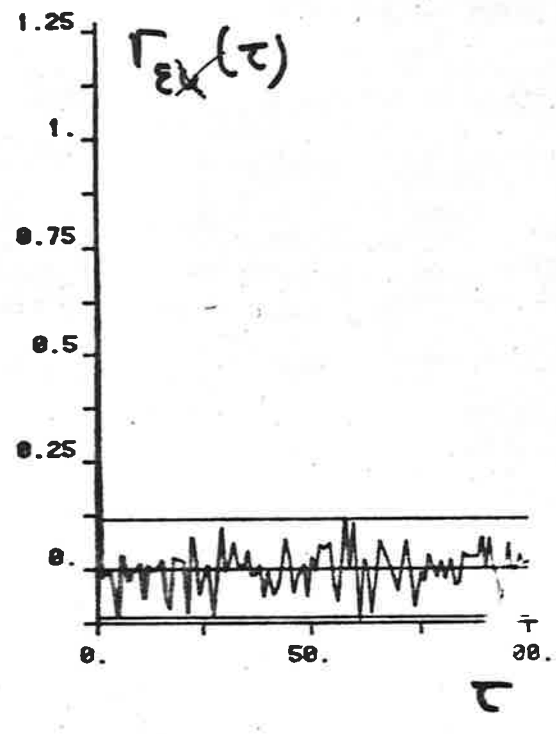
MEASURED DATA



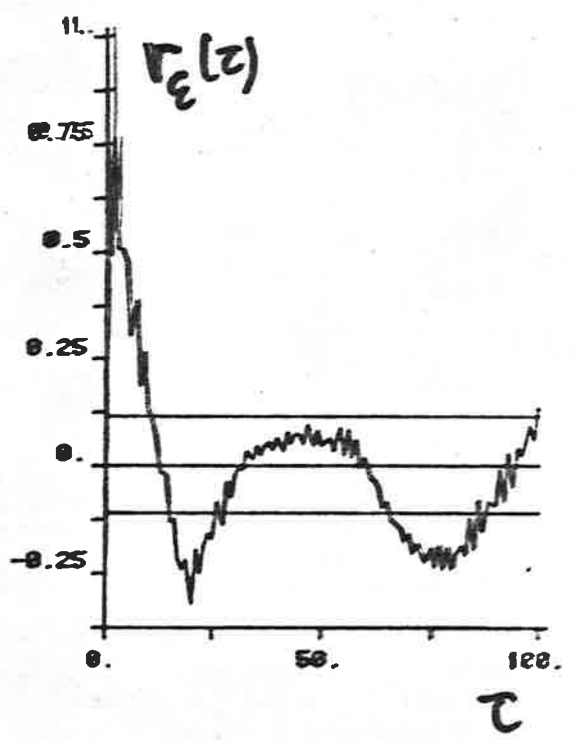
CALCULATION OF RESIDUALS & CORRELATIONS

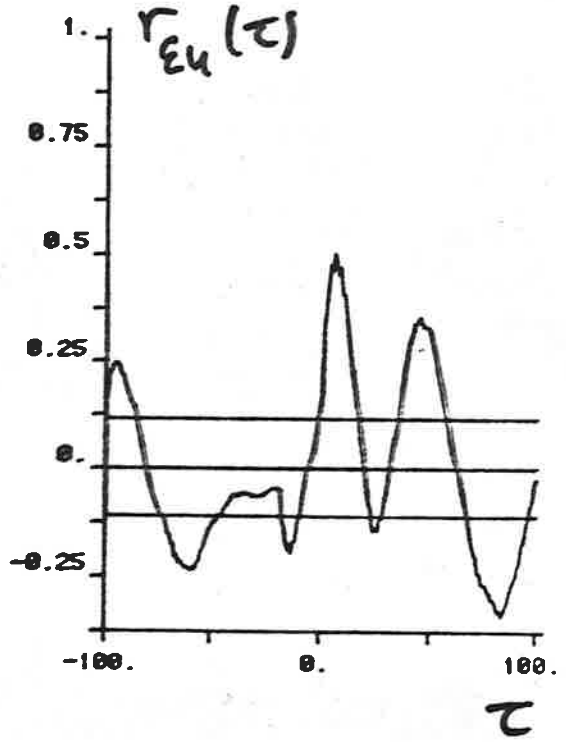
RESID E3 ← P3 WRK 100





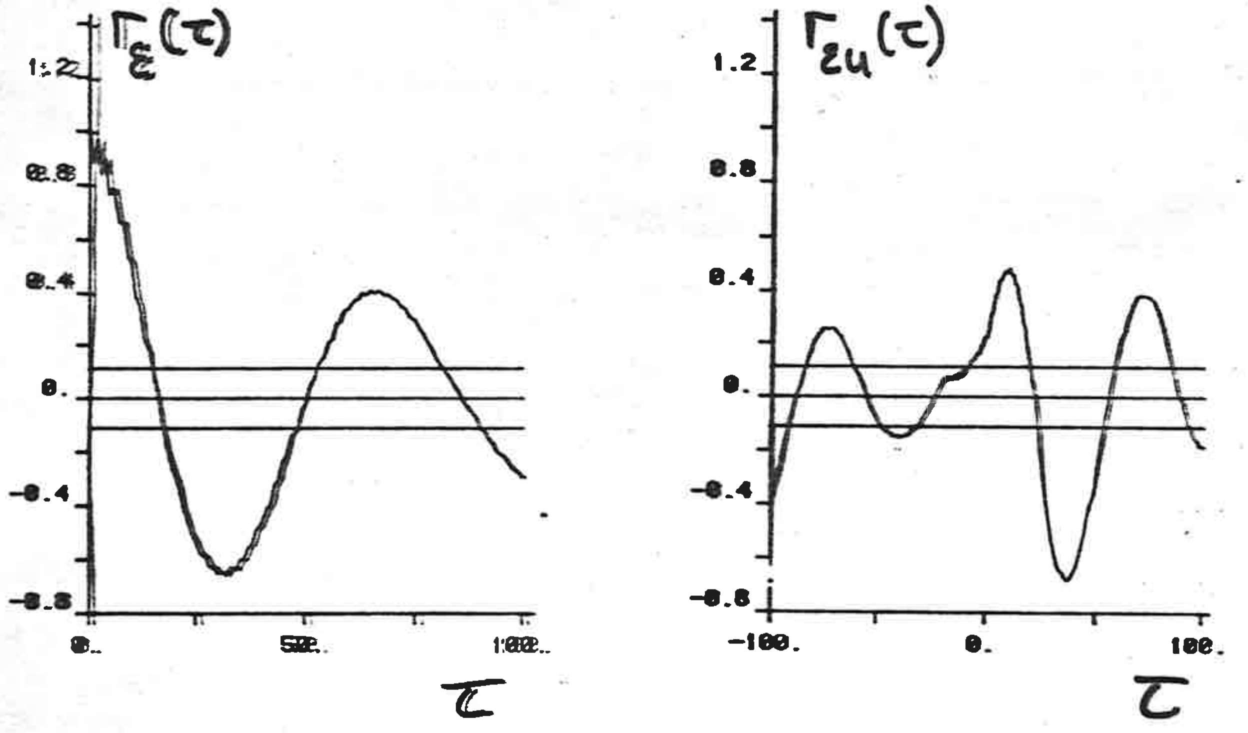
RESID E2 ← P2 WRK 100

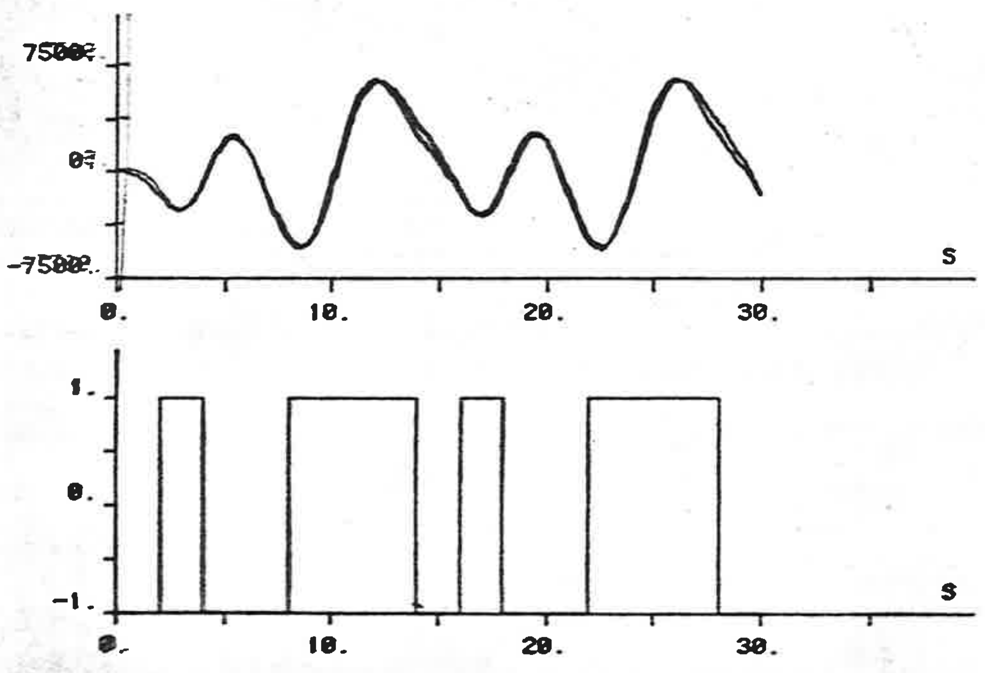




EXAMPLE OF MODEL VALIDATION

RESID E1 ← P1 WRK 100





TIME & COST

300 INPUT/OUTPUT PAIRS

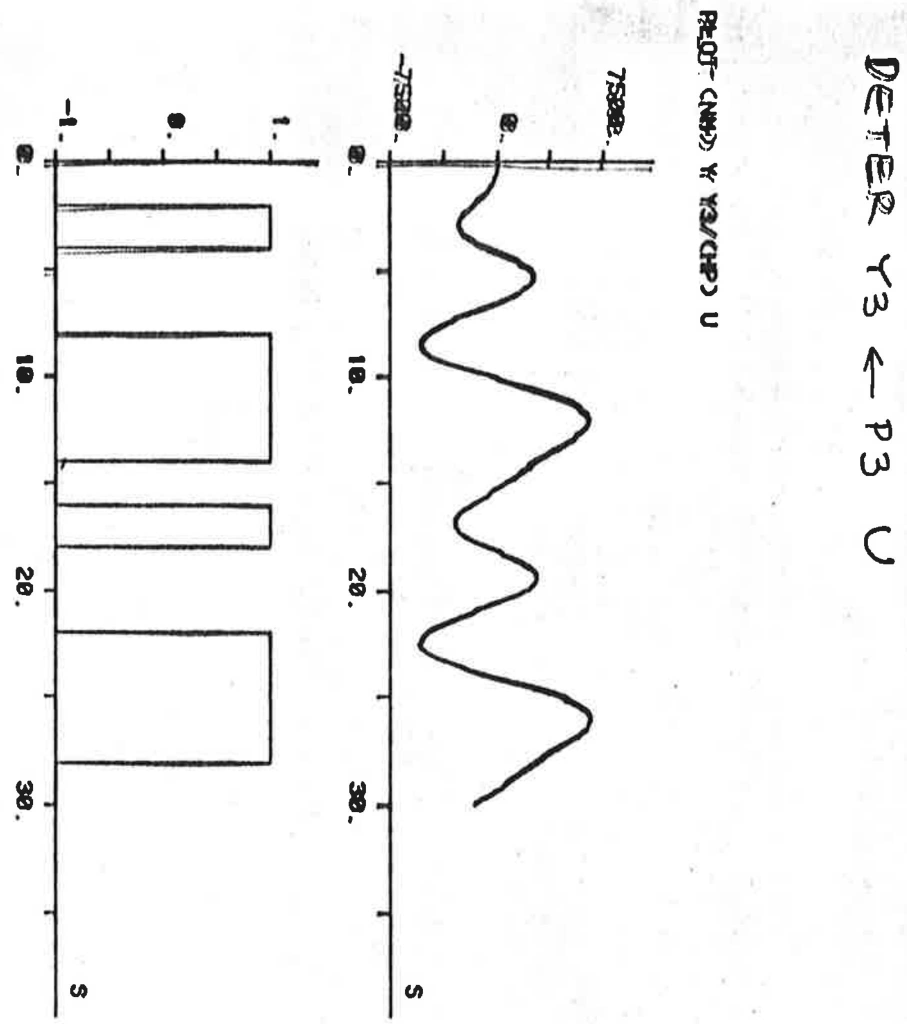
GENERATION OF DATA

PARAMETER ESTIMATION

PLOTTING OF CURVES

1.5 h TIME AT TERMINAL

TOTAL COST 82 Skr ≈ 27 YUAN



SPECIAL FEATURES OF ML & PRED. ERR.

- GREAT FLEXIBILITY WRT MODEL STRUCTURE
- DISTURBANCES ARE MODELED
- GREAT FLEXIBILITY WRT PARAMETRIZATION.
  "PHYSICAL" PARAMETERS & CONTINUOUS TIME MODELS CAN BE USED
- THEORETICALLY REASONABLE WELL UNDERSTOOD
- WILL OFTEN REQUIRE SUBSTANTIAL CALCULATIONS

PRACTICAL EXPERIENCES

PAPER MACHINES
DRUM BOILERS
DISTILLATION COLUMNS
NUCLEAR REACTORS
ACTIVATED SLUDGE PROCESSES
SHIP STEERING DYNAMICS
THERMAL HEAT CONDUCTION
MACROECONOMICS
PHARMACOKINETICS
INSULIN KINETICS

WHERE DOES ML & PE FIT INTO THE MODELING WORK?

- EXPLORATORY PHASE ASSUME A CANONICAL MISO MODEL. FIT TO DATA AND TEST?
- FINAL PARAMETER ESTIMATION PHASE. ASSUME PHYSICAL MODEL WITH ALL AVAILABLE INFORMATION. FIT PARAMETERS AND VALIDATE?

STOCHASTIC CONTROL THEORY AND APPLICATIONS

1. INTRODUCTION
2. LINEAR THEORY
   - INPUT - OUTPUT MODELS
   - STATE MODELS
3. MODELING & IDENTIFICATION
4. AN APPLICATION
5. NONLINEAR THEORY
6. SELF-TUNING CONTROL
7. AN APPLICATION
8. CONCLUSIONS

EXPERIENCES OF SYSTEM IDENTIFICATION APPLIED TO SHIP STEERING DYNAMICS

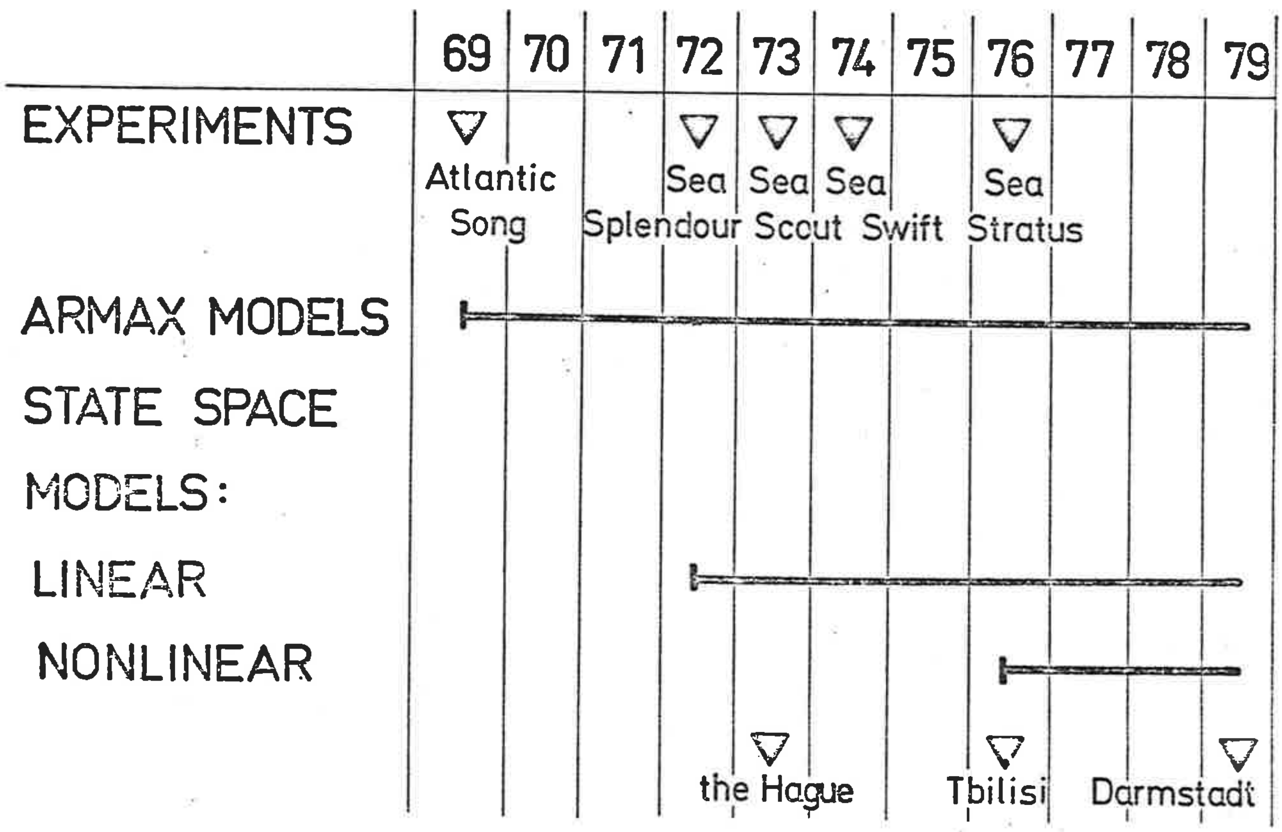
C.G. KÄLLSTRÖM
SWEDISH STATE SHIPBUILDING
EXPERIMENTAL TANK
(SSPA)

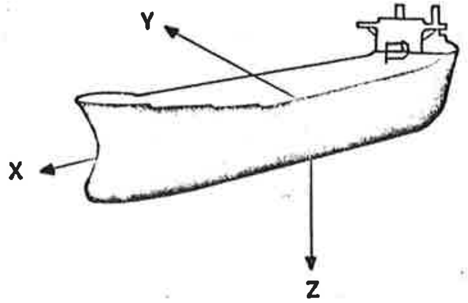
K.J. ÅSTRÖM
LUND INSTITUTE OF
TECHNOLOGY
(LTH)

1 ulan pil
5 med pil

1. INTRODUCTION
2. SHIP STEERING DYNAMICS
3. IDENTIFICATION METHODS
4. EXPERIMENTS
5. INTERESTING OBSERVATIONS
6. CONCLUSIONS

- ARE IDENTIFICATION METHODS OF ANY USE FOR DETERMINING SHIP STEERING DYNAMICS ?
- HOW CAN IDENTIFICATION METHODS BE USED IN AN EXERCISE IN DYNAMICAL MODELING ?
- WHAT IDENTIFICATION METHODS ARE APPROPRIATE ?





SIX DEGREES OF FREEDOM:

SURGE, SWAY, HEAVE

ROLL, PITCH, YAW

SENSOR AND ACTUATOR DYNAMICS

DISTURBANCES

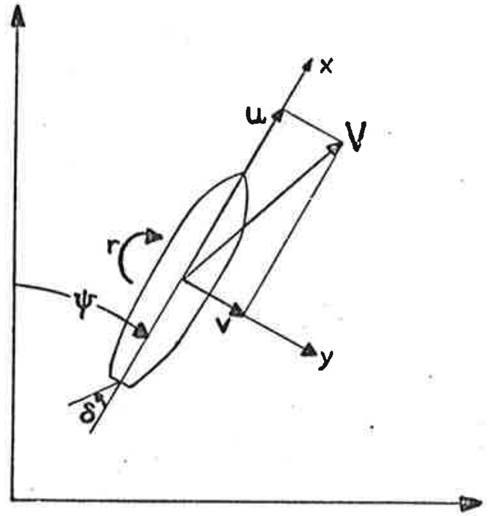
WIND, WAVES, CURRENTS

DIFFICULTIES

SIMPLIFICATIONS

SEPARATION

NORMALIZATION



EQUATIONS OF MOTION:

$$m(\dot{u} - vr - x_G r^2) = X + X_{disturbance}$$

$$m(\dot{v} + ur + x_G \dot{r}) = Y + Y_{disturbance}$$

$$I_z \dot{r} + m x_G(\dot{v} + ur) = N + N_{disturbance}$$

THE HYDRODYNAMIC FORCES X AND Y AND THE TORQUE N ARE FUNCTIONS OF THE MOTION, E.G.:

$$X = X(u, v, r, \delta, \dot{u}, \dot{v}, \dot{r})$$

EQUATIONS OF MOTION:

$$\begin{bmatrix} m - Y_{\dot{v}} & m x_G - Y_{\dot{r}} \\ m x_G - N_{\dot{v}} & I_z - N_{\dot{r}} \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} Y_v & Y_r - m u_0 \\ N_v & N_r - m x_G u_0 \end{bmatrix} \begin{bmatrix} v \\ r \end{bmatrix} + \begin{bmatrix} Y_\delta \\ N_\delta \end{bmatrix} \delta$$

STATE EQUATIONS:

$$\begin{bmatrix} \dot{v} \\ \dot{r} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ r \\ \psi \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{21} \\ 0 \end{bmatrix} \delta$$

TRANSFER FUNCTIONS:

$$G_{\psi\delta}(s) = K \frac{(s + 1/T_3)}{s(s + 1/T_1)(s + 1/T_2)} \approx$$

$$\approx \frac{K}{s(s + 1/T)} \quad \text{(Nomoto's model)}$$

$$G_{v\delta}(s) = K_v \frac{(s + 1/T_{3v})}{(s + 1/T_1)(s + 1/T_2)}$$

MODEL STRUCTURES:

DISCRETE TIME INPUT - OUTPUT (ARMAX) MODELS [IDPAC]:

$$y(t) + a_1 y(t-1) +\dots+ a_n y(t-n) = b_1 u(t-k-1) +\dots+ b_n u(t-k-n) + \lambda[e(t) + c_1 e(t-1) +\dots+ c_n e(t-n)]$$

CONTINUOUS TIME STRUCTURAL MODELS [LISPID]:

$$\begin{bmatrix} \dot{v} \\ \dot{r} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ r \\ \psi \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{21} \\ 0 \end{bmatrix} \delta$$

CRITERIA:

OUTPUT ERROR
MAXIMUM LIKELIHOOD
PREDICTION ERROR

$$\sum_i \sum_k y^T(t_i + \tau_k)\, Q\, y(t_i + \tau_k)$$

PARAMETER ESTIMATION

PROBLEM:

CONSIDER THE MODEL

$$dx = Ax\, dt + Bu\, dt + dw$$

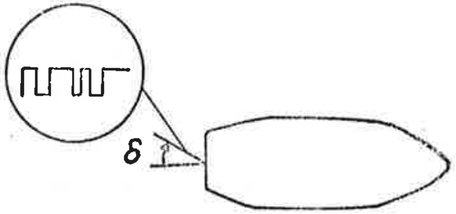
WHERE THE OUTPUT

$$y(t_k) = Cx(t_k) + Du(t_k) + e(t_k)$$

IS MEASURED AT DISCRETE TIMES

$$t_k, \quad k = 0, 1, \dots, N$$

DETERMINE THE PARAMETERS OF THE MODEL!



HEADING ψ
YAW RATE r
SWAY VELOCITY v

SIMPLE EQUIPMENT:

COMPASS .
MANUAL EXCITATION AND LOGGING

COMPLEX EQUIPMENT:

COMPASS, RATE GYRO, DOPPLER SONAR,
INERTIAL NAVIGATION
COMPUTER - CONTROLLED EXCITATION
AND LOGGING

25 EXPERIMENT CORRESPONDING TO
18 HOURS PERFORMED WITH
12 SHIPS HAVE BEEN ANALYSED

TANKERS:	SEA SPLENDOUR SEA SCOUT SEA SWIFT SEA STRATUS SEA SCAPE AK FERNSTRÖM NORSEMAN THORSHAMMER
CARGO SHIPS:	ATLANTIC SONG COMPASS ISLAND
FERRY:	BORE I
NAVAL CRAFT:	HIGH-SPEED PATROL BOAT

7 FREE-SAILING TESTS WITH
2 SCALE MODELS HAVE ALSO BEEN
ANALYSED