ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 6/7 — 1987

State-Space and Distributed Lag Modelling of Dynamic Economic Processes Based on Singular Value Decompositions (with an Application to the Dutch Economy)

Pieter W. OTTER, René VAN DAL *

ABSTRACT. – In this paper concepts and techniques from system theory will be used to obtain state-space (Markovian) models of dynamic economic processes instead of the usual ARMA-models. In this respect the concept of state will be reviewed and also Hankel factorisation, Hankel norm approximation and balanced realizations for stochastic models. The suggested procedures, which includes singular value decompositions, steady-state Kalman filtering and prediction-error estimation will be applied to some behavioral equations of a small macro-economic model representing the Dutch economy.

Modélisation espace d'états et retards échelonnés de processus économiques dynamiques fondée sur la décomposition en valeurs singulières, avec une application à l'économie hollandaise

RÉSUMÉ. – Dans cet article des concepts et des techniques de la théorie des systèmes sont utilisés pour obtenir des modèles « espace-d'états » (markovien) des processus économiques dynamiques au lieu des modèles ARMA habituels. Le concept d'état est discuté ainsi que la factorisation d'Hankel, l'approximation de la norme d'Hankel et les réalisations équilibrées des modèles aléatoires. Les procédures suggérées, qui incluent les décompositions en valeurs singulières, le filtre de Kalman invariant et la décomposition de l'erreur de prévision, sont appliquées à quelques équations de comportement d'un petit modèle macroéconomique représentant l'économie hollandaise.

* P. W. OTTER, R. VAN DAL: Econometrics Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, the Netherlands.

1 Introduction

There seems to be a growing interest among econometricians for the so-called state space model and its possibilities to model dynamic economic processes, see e. g. AOKI [1983, 1987], OTTER [1985, 1986] and WATSON and ENGLE [1983, 1985]. WATSON and ENGLE [1985] stated that "the state space model which underlies the Kalman filter should be of great interest to econometricians". Although we agree with their conclusion it is felt that the *concept* of state, as introduced by Kalman, is not fully appreciated by econometricians, a point already noticed by AOKI [1983]. Aim of the paper is to emphasize the concept of state in more detail by discussing concepts and results from system theory and apply them to (stochastic) econometric models.

The paper is organized as follows. In section two the concept of state will be discussed together with relevant concepts and results from system theory which will be used throughout the paper. In the same section population state space models for infinite multivariate distributed lag models with non-stochastic and stochastic inputs (exogenous variables) will be obtained together with approximations of the state space models. In section three sample state space models from finite Hankel matrices will be discussed together with some parameter estimation aspects. In section four finally the procedures will be applied to some behavioral equations of a small macro-economic model representing the Dutch economy.

2 The Concept of State, Balanced Realization and Hankel Norm Approximation

Loosely speaking we have that the state is a memory function which accumulates the information from the past behavior of the system in as much it is relevant for the future behavior of the system. The concept of state has its roots in physics and can be seen as an abstraction of the state of e. g. a particle, which is determined by its position and moment. The future motion of a particle is solely determined by its state (and external forces); how the (present) state has been reached is unimportant for the future behavior of the particle. In the sequel we discuss some econometric examples in which the concept of state is being illustrated.

Consider the following scalar infinite distributed lag model without the disturbances, i. e.

(1) $$y_t = \sum_{i=0}^{\infty} g_i u_{t-i}$$

with $y_t \in \mathbb{R}$ as endogenous (output) variable and $u_t \in \mathbb{R}$ as exogenous (input) variable. By imposing certain patterns for the coefficients (impulse response functions) $\{g_i\}$ the model can be simplified in order to estimate the coefficients. In the well-known Koyck model for instance we have as weighting scheme

$$g_i = \beta(1-\lambda)\lambda^i, \qquad 0<\lambda<1$$

see e. g. MADDALA [1977], which leads to the simplified model

(2) $$y_t = \lambda y_{t-1} + \beta(1-\lambda) u_t \qquad \text{or} \qquad y_{t+1} = \lambda y_t + \beta(1-\lambda) u_{t+1}$$

Comparing equation (2) with equation (1) we have a simplification and efficient description by equation (2) in the sense that in equation (2) one may forget *all* past inputs up to time t because the *future* behaviour of the model (y_{t+l}, $l \geq 1$) is solely determined by the *state* y_t and future input u_{t+l}, $l \geq 1$. In equation (1) the future behavior of the endogenous variable y depends on *all* past input values. Thus instead of remembering all past inputs we can in the Koyck model accumulate this information in the state y_t, which determines the future behavior. In systems terms we have that y_t is the state for equation (2). Loosely speaking we have that the state is the memory function which accumulates the information from the past (inputs) in as much it is relevant for the future behavior of the system. In the above example we have that, by imposing a certain weighting pattern for the coefficients, all information in the past inputs may be collected in the state y_t.

As a second example consider the familiar regression equation

$$y_t = x_t' \beta + \varepsilon_t, \qquad t = 1, 2, \ldots$$

with $y_t \in \mathbb{R}$ as endogenous variable; $x_t \in \mathbb{R}^q$ a vector of exogenous variables and $\varepsilon_t \sim \mathrm{NID}(0, \sigma_\varepsilon^2)$, to be referred to as Gaussian white-noise. The least-squares estimator for β based on $T \geq q$ observations $\{y_t, x_t, t=1, \ldots, T\}$ is given by

(3) $$b_T = (X_T' X_T)^{-1} \sum_{i=1}^{T} x_i y_i$$

with $X_T = (x_1, \ldots, x_T)'$. However, as well known, the least-squares estimator in recursive form can be written as

(4) $$b_T = b_{T-1} + K_T r_T$$

where K_T is a weighting (gain) matrix and $r_T = y_T - x_T' b_{T-1}$ is the prediction error with zero expectation, see e. g. BROWN *et al.* [1975] or OTTER [1978]. Comparing equation (3) and equation (4) it is seen that in equation (3) all past information $\{y_t, x_t, t=1, \ldots T-1\}$ is relevant to compute b_T, whereas in equation (4) this information is accumulated in the state b_{T-1}, or to say it in statistical terms, b_{T-1} is a *sufficient* statistic and we may "forget" the information in $\{y_t, x_t, t=1, \ldots, T-1\}$. Only the present value y_T, x_T is relevant to update b_{T-1}, which leads to the conclusion that equation (4) is a more efficient description of the least-squares estimator. Generalizing the agove procedure leads to the Kalman filter for regression models, see e. g. OTTER [1978] or WATSON and ENGLE [1983].

Kalman formalized the idea of state for dynamical systems, not be imposing certain weighting patterns for the coefficients as usually is being done in econometrics but by introducing an auxiliary space, the so-called state space, which leads to a state-space or Markovian representation of dynamic processes instead of the familiar ARMA (X)-representation. In the sequel some essential results will be shortly reviewed with respect to state space modelling.

2.1. Models with Non-Stochastic Inputs

Consider again the infinite distributed lag equation (1) but now for the multivariate case without disturbances, i. e.,

$$y_t = \sum_{i=0}^{\infty} G_i u_{t-i} = G(L)\, u_t \tag{5}$$

with output $y_t = \mathbb{R}^p$ and input $u_t \in \mathbb{R}^q$ where $p \leqslant q$. The matrix polynomial, to be referred to as transfer function or impulse response function, is given by $G(L) = G_0 + G_1 L + G_2 L^2 + \ldots$ where L denotes the usual lag operator. The future behavior (at time t) of the system, namely $\{y_{t+i}, i=1, 2, \ldots\}$ is a linear function of the *past* inputs $\{u_t, u_{t-1}, \ldots\}$ and future inputs $\{u_{t+1}, u_{t+2}, \ldots\}$ i. e.

$$\begin{bmatrix} y_{t+1} \\ y_{t+2} \\ \vdots \end{bmatrix} = \begin{bmatrix} G_1 & G_2 & G_3 \\ G_2 & G_3 & \ldots \\ G_3 & \ldots & \\ \vdots & & \end{bmatrix} \times \begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \end{bmatrix} + \begin{bmatrix} G_0 & 0 & \ldots \\ G_1 & G_0 & \ldots \\ \vdots & \vdots & \end{bmatrix} \times \begin{bmatrix} u_{t+1} \\ u_{t+2} \\ \vdots \end{bmatrix}$$

or

$$y_t^+ = H u_t^- + T u_t^+$$

where H is the so-called infinite Hankel matrix and T a Toeplitz matrix. If the rank of the Hankel matrix defined by

$$r(H) = \sup_{N, N'} r(H_{N,N'}) = \lim_{N, N' \to \infty} r(H_{N,N'})$$

is finite dimensional. say $r(H) = n < \infty$ then it can be proven see e. g.

WILLEMS [1984] or KAILATH [1980] that equation (4) admits the following so-called state space realization (model) denoted by $\Sigma(A, B, C, D)$:

$$x_{t+1} = A\, x_t + B\, u_t \quad \text{initial condition } x_0 = 0$$
$$y_t = C\, x_t + D\, u_t$$

with $D = G_0$ and $G_i = CA^{i-1}B$. The state vector is given by x_t with $\dim(x_t) = n$. The finite Hankel matrix $H_{N, N'}$ is defined as

$$H_{N, N'} = \begin{bmatrix} G_1 & \dots & G_{N'} \\ \vdots & & \\ G_N & \dots & G_{N+N'-1} \end{bmatrix}$$

It is easily seen that the Hankel matrix H with $r(H) = n$ can be factorized as

$$H = QP = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} (B\, AB\, A^2B \dots)$$

where Q denotes in system theory the (extended) observability matrix and where P denotes the (extended) reachability matrix. If $r(Q) = n$ then it is said that the pair (A, C) is observable; if $r(P) = n$ then it is said the pair (A, B) is reachable. See for a discussion of the observability matrix Q and reachbility matrix P in terms of parameter-identifiability OTTER [1986].

Consider a (coordinate) transformation of the state-vector x_t, i. e. $\tilde{x}_t = T\, x_t$ with $\det(T) \neq 0$. The new state-space model is $\Sigma(\tilde{A}, \tilde{B}, \tilde{C}, D)$, i. e.

$$\tilde{x}_{t+1} = \tilde{A}\, \tilde{x}_t + \tilde{B}\, u_t$$
$$y_t = C\, \tilde{x}_t + D\, u_t$$

with $\tilde{A} = TAT^{-1}$, $\tilde{B} = TB$ and $\tilde{C} = CT^{-1}$.

The transfer function of the new state-space model is, after eliminating the state-vector $\tilde{x}_t$, given by

$$\tilde{G}(L) = [CT^{-1}(IL^{-1} - TAT^{-1})^{-1}TB + D]\, u_t$$
$$= [C(IL^{-1} - A)^{-1}B + D]\, u_t = G(L)$$

where $G(L) = \sum_{i=0}^{\infty} G_i L^i$ with $G_0 = D$ and $G_i = CA^{i-1}B$, $i \geqslant 1$ is the transfer function of Σ (A, B, C, D). Because the transfer functions are equal in both models, it is said that the state-space models are T-equivalent. This non-uniqueness has led to the development of the so-called canonical form, see e. g. PICCI [1982]. In this paper we discuss another particular state-space model (realization), the so-called (internally) balanced model, see e. g. MOORE [1981], because this type of realization is suitable for approximation to be discussed in the sequel. Consider again the realization $\Sigma(A, B, C, D)$

in which the pair (A, B) is reachable and the pair (A, C) observable with $\dim(x_t) = n$. The so-called reachbility Gramian W is given by

$$\min_{0 \xrightarrow{u} x_0} \sum_{t=-\infty}^{-1} \|u_t\|^2 = x_0' \mathrm{W}^{-1} x_0$$

where $\mathrm{W} = \sum_{t=0}^{\infty} \mathrm{A}^t \mathrm{BB}'(\mathrm{A}')^t$ is the unique solution of the Lyapunov equation $\mathrm{W} - \mathrm{AWA}' = \mathrm{BB}'$, where $\|.\|$ denotes the Euclidean norm. The interpretation of $x_0' \mathrm{W}^{-1} x_0$ is that it represents the minimum amount of "energy" measured by the "variance" $\|u_t\|^2$ of the input vector required to drive the state vector from zero to x_0 at $t = 0$.

Thus $x_0' \mathrm{W}^{-1} x_0$ "large" means that the state vector is difficult to move towards x_0 from zero compared with $x_0' \mathrm{W}^{-1} x_0$ "small". The so-called observability Gramian M is given by

$$\sum_{t=0}^{\infty} \|y_t\|^2 = x_0' \mathrm{M} x_0$$

where $\mathrm{M} = \sum_{t=0}^{\infty} (\mathrm{A}')^t \mathrm{C}'\mathrm{CA}^t$ is the unique solution of the Lyapunov equation $\mathrm{M} - \mathrm{A}'\mathrm{MA} = \mathrm{C}'\mathrm{C}$ with $\mathrm{M} = \mathrm{M}' > 0$. The intuitive interpretation of $x_0' \mathrm{M} x_0$ is that it gives the amount of observation fluctuations measured by the "variance" $\|y_t\|^2$ available in the initial state x_0. Thus $x_0' \mathrm{M} x_0$ "large" means that x_0 is "easy" to observe compared with $x_0' \mathrm{M} x_0$ "small".

DEFINITION : The realization $\Sigma(\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D})$ is said to be (internally) balanced if $\mathrm{W} = \mathrm{M}$.

The reason of balancing is to neglect components of the state vector which are difficult to move and which are difficult to observe (they contribute little to the observation "variance"). However a component of the state vector may be difficult to move but easy to observe or easy to move and difficult to observe. Therefore balancing is a compromise between the two, see for details WILLEMS [1984].

Consider Aoki's example (AOKI [1983] p. 69) where the state-space model if given by

$$\begin{bmatrix} x_{1,t+1} \\ x_{2,t+1} \end{bmatrix} = \begin{bmatrix} -1/2 & 0 \\ 0 & -1/3 \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + \begin{bmatrix} 10^{-6} \\ 10^{6} \end{bmatrix} u_t$$

$$y_t = [10^6 \quad 10^{-6}] \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix}$$

with impulse response functions $g_i = \mathrm{CA}^{i-1}\mathrm{B} = (-1/2)^{i-1} + (-1/3)^{i-1}$, $i \geqslant 1$. From the state-space model it can be seen that, whenever there is a change in u_t, the second component of the state vector is enormously magnified, but hardly observed by the output variable y. The opposite is true for the first component of the state vector. By solving the Lyapunov equation

$W - AWA' = BB'$ we have as reachability Gramian

$$W = \begin{bmatrix} 4/3.10^{-12} & 6/5 \\ 6/5 & 9/8.10^{12} \end{bmatrix}$$

and by solving the equation $M - AMA' = CC'$ we have as observability Gramian

$$M = \begin{bmatrix} 4/3.10^{12} & 6/5 \\ 6/5 & 9/8.10^{-12} \end{bmatrix}$$

and it is seen that the Gramians are hardly in "balance" ($W \neq M$). Rescaling the components of the state-vector by the transformation

$$\tilde{x}_t = T x_t$$

where $x_t = (x_{1,t} x_{2,t})'$ and

$$T = \begin{bmatrix} 10^6 & 0 \\ 0 & 10^{-6} \end{bmatrix}$$

yields as state-space model

$$\begin{bmatrix} \tilde{x}_{1,t+1} \\ \tilde{x}_{2,t+1} \end{bmatrix} = \begin{bmatrix} -1/2 & 0 \\ 0 & -1/3 \end{bmatrix} \begin{bmatrix} \tilde{x}_{1,t} \\ \tilde{x}_{2,t} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_t$$

$$y_t = (1 \quad 1) \begin{bmatrix} \tilde{x}_{1,t} \\ \tilde{x}_{2,t} \end{bmatrix}$$

with as reachability and observability Gramians

$$W = M = \begin{bmatrix} 4/3 & 6/5 \\ 6/5 & 9/8 \end{bmatrix}$$

and hence this model is, by rescaling, internally balanced.

A balanced realization for the infinite distributed lag model can be obtained as follows. The Hankel matrix can be factorized as $H = QP$ with $Q' = (C'(CA')(CA^2)' \ldots)$ as the observability matrix and $P = (B A\text{-} B A^2 B \ldots)$ as the reachability matrix. Following Kung and Lin [1981], the Hankel matrix H with $r(H) = n$ can be written as the singular value decomposition (SVD)

$$H = \bar{U} \bar{\Sigma} \bar{V}' = \sum_{i=1}^{n} \sigma_i u_i v_i'$$

where $\bar{\Sigma} = \text{diag}(\sigma_1, \ldots, \sigma_n)$ with singular values $\sigma_1 \geqslant \ldots \geqslant \sigma_n > 0$; u_i and v_i are infinite singular vectors with $u_i' u_j = v_i' v_j = 0$ if $i \neq j$ and 1 if $i = j$; $\bar{U} = (u_1, \ldots, u_n)$ and $\bar{V} = (v_1, \ldots, v_n)$.

Let $Q = \bar{U} \bar{\Sigma}^{1/2}$ and $P = \bar{\Sigma}^{1/2} \bar{V}'$ then $H = QP = \bar{U} \bar{\Sigma} \bar{V}'$, Denote by "$\uparrow$" the shift-up operator defined by

$$QA = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} A = \begin{bmatrix} CA \\ CA^2 \\ \vdots \end{bmatrix} = Q^{\uparrow}$$

The balanced realization (state space model) $\Sigma(A_b, B_b, C_b, D)$ is given by $A_b = Q^+ Q^\uparrow$; B_b = the first q columns of P (q is the dimension of the input vector); C_b = first p rows of Q (p is the dimension of the output vector) and $D = G_0$. Here Q^+ denotes a pseudo-inverse e. g. $(Q'Q)^{-1}Q'$. The observability Gramian

$$M = Q'Q = (C'(CA)' \dots)(C'(CA)' \dots)$$
$$= \sum_{t=0}^{\infty} (A)^t C' CA^t = \bar{\Sigma}^{1/2} \bar{U}' \bar{U} \bar{\Sigma}^{1/2} = \bar{\Sigma}$$

and the controllability Gramian

$$W = PP' = (B\, AB\, A^2 B \dots)(B'(AB)' \dots)'$$
$$= \sum_{t=0}^{\infty} A^t BB'(A')^t = \bar{\Sigma}^{1/2} \bar{V}' \bar{V} \bar{\Sigma}^{1/2} = \bar{\Sigma}$$

and hence the realization is balanced.

Summarizing the foregoing but now for the infinite distributed lag model with a disturbance vector $\delta_t \sim NID(0, V_\delta)$ i. e.

(6) $$y_t = G(L)\, u_t + \delta_t = \bar{y}_t + \delta_t$$

where $\bar{y}_t$ is the "systematic" part of y_t, we have for the systematic part $\bar{y}_t$ a balanced realization $\Sigma(A_b, B_b, C_b, D)$ provided that $r(H) = n$. For equation (6) we have the stochastic state space realization

$$x_{t+1} = A_b x_t + B_b u_t \qquad \text{initial condition } x_0 = 0$$
$$y_t = C_b x_t + D\, u_t + \delta_t$$

where the dimension of the state vector x_t equals n. Because the measurements $\{y_t\}$ are noisy the sequence $\{x_t\}$ can be estimated optimally from the input and output sequence $\{y_t, u_t\}$ by using the Kalman filter, see e. g. SAGE and MELSA [1971]. According to BERTSEKAS [1976] however we have that, because the pair (A_b, C_b) is observable and (A_b, B_b) is reachable "in the large" the Kalman filter reaches its steady state given by the following equations

$$\hat{x}_{t+1\,|\,t} = A_b \hat{x}_{t\,|\,t-1} + B_b u_t + A_b K e_t$$

where $K = \Sigma C'(C'\Sigma C + V_\delta)^{-1}$ is the steady state Kalman gain and Σ is the steady state *a priori* error covariance matrix of the Kalman MSE prediction $\hat{x}_{t+1\,|\,t}$ for all t.

Σ is the positive definite solution of the algebraic Riccati equation

$$\Sigma = A_b [\Sigma - \Sigma C_b' (C_b \Sigma C_b' + V_\delta)^{-1} C_b \Sigma] A_b'.$$

The prediction error is given by

$$e_t = y_t - \hat{y}_{t\,|\,t-1} = y_t - C_b \hat{x}_{t\,|\,t-1} - D\, u_t$$

which is a Gaussian white noise process with covariance $V_e = C_b \Sigma C_b' + V_\delta$, see e. g. BERTSEKAS [1976].

For the stochastic state space model we have "in the large" the steady state Kalman filter, which itself is a stochastic state space model, i. e.

$$\hat{x}_{t+1\,|\,t} = A_b \hat{x}_{t\,|\,t-1} + B_b u_t + A_b K e_t$$

initial condition (in the large) $\hat{x}_0 = \mu$.

$$y_t = C_b \hat{x}_{t\,|\,t-1} + D u_t + e_t$$

which will be denoted by $\Sigma_{KF}(A_b, B_b, C_b, D)$.

2.2. Models with Stochastic Inputs

Suppose that the input process $\{u_t\}$ in the infinite distributed lag model

$$y_t = G(L)\, u_t + \delta_t$$

is a stationary, stochastic process which admits an ARMA (p, q) representation, i. e.

$$A(L)\, u_t = B(L)\, \eta_t$$

where it is assumed that $\det(A(L)) \neq 0$ and $\{\eta_t\}$ is a Gaussian white noise process with covariance V_η of full rank which can be factorized as $V_\eta = F_\eta F'_\eta$ where F_η^{-1} exists.

The matrix polynomials $A(L)$ and $B(L)$ are given by

$$A(L) = I + A_1 L + A_2 L^2 + \ldots + A_p L^p$$

and

$$B(L) = B_0 + B_1 L + \ldots + B_q L^q$$

Substituting

$$u_t = A^{-1}(L)\, B(L)\, \eta_t = \theta(L)\, \eta_t$$

into the infinite lag model yields

$$y_t = G(L)\, \theta(L)\, \eta_t + \delta_t = \psi(L)\, \tilde{\eta}_t + \delta_t$$

where $\tilde{\eta}_t = F_n^{-1} \eta_t$ is standard Gaussian white noise ($\tilde{\eta}_t \sim NID(0, I)$) and $\psi(L) = G(L)\, \theta(L)\, F_\eta$.

It is assumed that $\psi(L)$ is stable. It is seen that $\{y_t\}$ is a stationary process with zero mean with as input the standard white noise process $\{\tilde{\eta}_t\}$. Relating the future outputs to the past inputs we have

$$\begin{bmatrix} y_{t+1} \\ y_{t+2} \\ \vdots \end{bmatrix} = \begin{bmatrix} \psi_1 & \psi_2 & \cdots \\ \psi_2 & \psi_3 & \cdots \\ \psi_3 & \cdots & \cdots \\ \vdots & \vdots & \end{bmatrix} \begin{bmatrix} \tilde{\eta}_t \\ \tilde{\eta}_{t-1} \\ \vdots \end{bmatrix} + \begin{bmatrix} \psi_0 & 0 & & \cdots \\ \psi_1 & \psi_0 & 0 & \cdots \\ \vdots & \vdots & & \end{bmatrix} \begin{bmatrix} \tilde{\eta}_{t+1} \\ \tilde{\eta}_{t+2} \\ \vdots \end{bmatrix} + \begin{bmatrix} \delta_{t+1} \\ \delta_{t+2} \\ \vdots \end{bmatrix}$$

or

$$y_t^+ = H_\psi \tilde{\eta}_t^- + T_\psi \tilde{\eta}_t^+ + \delta_t^+$$

The *covariance* of the future outputs y_t^+ and the past inputs $\tilde{\eta}_t^-$ is given by the Hankel matrix H_ψ, i. e.

$$E\{y_t^+ (\tilde{\eta}_t^-)'\} = H_\psi E\{\tilde{\eta}_t^- (\tilde{\eta}_t^-)'\} + T_\psi E\{\tilde{\eta}_t^+ (\tilde{\eta}_t^-)'\} + (E\{\delta_t^+ (\tilde{\eta}_t^-)'\} = H_\psi$$

Here $E\{.\}$ denotes the expectation operator. If $r(H_\psi) = n$ then we can apply the balanced realization procedure given before which gives the state space model

$$x_{t+1} = A_b x_t + B_b \tilde{\eta}_t \qquad \text{initial condition } x_0 = 0.$$
$$y_t = C_b x_t + D \tilde{\eta}_t + \delta_t$$

Because y_t is stochastic we may associate the steady state Kalman filter to the above model, that is,

$$\hat{x}_{t+1|t} = A_b \hat{x}_{t|t-1} + B_b \tilde{\eta}_t + A_b K e_t \qquad \text{initial condition:}$$
$$y_t = C_b \hat{x}_{t|t-1} + D \tilde{\eta}_t + e_t, \qquad \hat{x}_0 = \mu.$$

In the next section it is seen that by identifying the input ARMA (p, q) model along the lines of BOX and JENKINS [1970], realizations of $\{\tilde{\eta}_t\}$ are given by the *input* prediction-errors. Thus the above steady state Kalman filter is a function of the *input* prediction errors (observations of $\tilde{\eta}_t$) and *output* prediction errors $\{e_t\}$.

2.3. Model Reduction and Hankel Norm Approximation

Suppose that the dimension of the state vector in the balanced realization is rather high and that some of the singular values are "close" to zero. The question arises whether the state space model with $\dim(x_t) = n$ may be approximated by a lower order state space model yet maintaining the input/output behavior "as much as possible". The problem of model reduction is that an approximation procedure must preserve the Hankel structure. Recently there has been a breakthrough, based on Hankel norm approximation, see e. g. GLOVER [1984] but in this paper we follow a simple procedure suggested by KUNG and LIN [1981]. Let

$$F(L) = C(L^{-1} I - A)^{-1} B$$

be the transfer function (without D) of the state space model $\Sigma(A, B, C, D)$ with $\dim(x_t) = n$ and where L^{-1} is the forward lag operator. The Hankel norm of $F(L)$ is defined as

$$\| F(L) \|_H = \bar{\sigma}(H)$$

where $\bar{\sigma}(H)$ is the greatest singular value of the Hankel matrix H of $\Sigma(A, B, C, D)$. The Hankel norm lies between the conventional L_2 and L_∞

norm. Given the Hankel matrix with rank n the problem is to find a Hankel matrix $\tilde{H}$ with *given* $r(\tilde{H})=k<n$ such that $\|H-\tilde{H}\|=\bar{\sigma}(H-\tilde{H})$ is minimum. The SVD of H given before is

$$H=\bar{U}\bar{\Sigma}\bar{V}'=\sum_{i}^{n}\sigma_i u_i v_i'$$

with $\bar{U}=(u_1, \ldots, u_n)$, $\bar{\Sigma}=\mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ and $\bar{V}(v_1, \ldots, v_n)$. An (non-unique) approximant L_R of H with $r(L_R)=k$ which minimizes the Hankel norm is

$$L_R=\sum_{i=1}^{k}\sigma_i u_i v_i'$$

Because of the minimax property of singular value decomposition we have

$$\min_{L_R:\, r(L_R)\leq k}\|H-L_R\|_H=\sigma_{k+1}$$

see e. g. WILLEMS [1984]. L_R is in general *not* a Hankel matrix and therefore does not represent a dynamic system of order k. However the following decomposition still holds

$$L_R=\bar{Q}\bar{P}$$

with $\bar{Q}=U_1\Sigma_{(k)}^{1/2}$ and $\bar{P}=\Sigma_{(k)}^{1/2}V_1'$, where $U_1=(u_1, \ldots, u_k)$ and $V_1=(v_1, \ldots, v_k)$ and $\Sigma_{(k)}=\mathrm{diag}(\sigma_1, \ldots, \sigma_k)$. Applying the balanced realization procedure given before leads to the so-called

Principal Hankel Component (PHC) Approximation:

$$A_{b,R}=\bar{Q}^{+}\bar{Q}^{\uparrow}=\Sigma_{(k)}^{-1/2}U_1'U_1^{\uparrow}\Sigma_{(k)}^{1/2};$$

$$B_{b,R}=\text{first } q \text{ columns of } \bar{P};$$

$$C_{b,R}=\text{first } p \text{ rows of } \bar{Q} \text{ and } D=G_0.$$

See for details KUNG and LIN [1981] and references therein. A simple approximation procedure (quasi-balancing) which is equivalent to the PHC Approximation procedure, see SILVERMAN and BETTAYEB [1980], is by considering a submodel $\Sigma(A_{11}, B_1, C_1, D)$ with dimension k of the state-vector of the balanced realization $\Sigma(A_b, B_b, C_b, D)$ with dimension of the state vector equal to n, where

$$A_b=\left[\begin{array}{c|c}A_{11} & A_{12}\\ \hline \underbrace{A_{21}}_{k\text{-columns}} & A_{22}\end{array}\right]\}\,k\text{-rows},\qquad B_b=\left[\begin{array}{c}B_1\\ \hline B_2\end{array}\right],\qquad C=(C_1|C_2)$$

In the sub-model only the first k components of the state vector of $\Sigma(A_b, B_b, C_b, D)$ are being considered. A possible model reduction for the steady state Kalman filter can be obtained in a similar way.

In section 2.2 we have seen that for the models with prewhitened inputs the Hankel matrix H_ψ is the covariance between the output and the prewhitened input. Singular value decomposition of H_ψ in this case is equivalent to the canonical correlation procedure; the singular values σ_i, $i=1, \ldots, n$ are the canonical correlation coefficients. The mutual information between y_t^+ and $\tilde{\eta}_t^-$ can be expressed in H_ψ and GELFAND and YAGLOM [1959] have shown that

$$I(y_t^+, \tilde{n}_t^-) = -\frac{1}{2} \sum_{i=1}^{n} \log(1-\sigma_i^2), \qquad \forall t$$

where $I(y_t^+, \tilde{\eta}_t^-)$ denotes the information about y_t^+ contained in $\tilde{\eta}_t^-$, the past input innovations. Following DESAI and PAL [1982] we have that a performance index for the k-th order approximation with $k<n$ of a state space realization with the dimension of the state vector equal to n is given by

$$\xi = \sum_{i=1}^{k} \log(1-\sigma_i^2) / \sum_{i=1}^{n} \log(1-\sigma_i^2)$$

where ξ represents the fraction of the mutual information retained in the k-th order approximation. For a 'good' approximation the value of ξ should be 'close' to one. For a test statistic based on canonical correlation coefficients we refer to section 3.

Summarizing this section we have the following. The infinite distributed lag model with *given* impulse response functions $\{G_i\}$ may be parsimoniously modelled by a balanced realization with dimension of the state vector equal to n if $r(H)=n$. If, say, the last $(n-k)$ singular values of the Hankel matrix are "close to zero" then a k-th order approximation is given by considering a reduced order model with the k first component of the state vector (quasi balancing). For the infinite distributed lag model with Gaussian noise δ_t but with known impulse response functions $\{G_i\}$ we have that the systematic part can be modelled as a balanced realization, whereas the balanced realization with the noise component δ_t is a stochastic state space model for which the (steady state) Kalman filter applies. The steady state filter is essentially a stochastic state space model in the so-called prediction error (innovations) form, see e. g. GOODWIN and PAYNE [1977], and also PICCI [1982]. A possible model reduction of the steady state filter may be obtained by considering only the k first components of the state vector (quasi-balancing).

3 Model Determination from a Sample

3.1. Finite Hankel Approximation

Until now we discussed balanced realizations for infinite distributed lag models with non-stochastic and stochastic inputs. It was assumed that the impulse response functions $\{G_i\}$ are known in advance and in case of stochastic inputs that the polynomial matrices $A(L)$ and $B(L)$ are known. Now we relax this assumption and discuss balanced realizations from finite data samples. We begin with the distributed lag model with non-stochastic inputs, i. e.

$$y_t = G(L)\, u_t + \delta_t$$

The problem is to estimate the impulse response functions from finite data samples. Several procedures exist, for instance Matrix Fraction Descriptions (MFD), see e. g. KAILATH [1980] or GOODWIN and PAYNE [1977]. Padé approximations, see KUNG and LIN [1981] and several econometric procedures, see e. g. MADDALA [1977]. As a first approach the following simple procedure is suggested. Decompose the distributed lag model as

$$y_t = \sum_{i=0}^{r} G_i u_{t-i} + \sum_{j=r+1}^{\infty} G_j u_{t-j} + \delta_t, \qquad t=0, \ldots, T$$

Assume r such that $\sum_{j=r+1}^{\infty} G_j u_{t-j} \approx C$ or at least slowly time-varying. The approximated regression model is

$$\begin{bmatrix} y'_0 \\ y'_1 \\ \vdots \\ y'_T \end{bmatrix} = \begin{bmatrix} u'_0 & u'_{-1} & \ldots & u'_{-r} & 1 \\ u'_1 & \ldots & & & 1 \\ \vdots & & & & \vdots \\ u'_T & \ldots & & u'_{T-r} & 1 \end{bmatrix} \times \begin{bmatrix} G'_0 \\ G'_1 \\ \vdots \\ G'_r \\ C' \end{bmatrix} + \begin{bmatrix} \delta_0 \\ \vdots \\ \delta_T \end{bmatrix}$$

from which least-squares estimates $\{\hat{G}_i, i=0, \ldots, r\}$ can be obtained. The *finite* $Np \times Nq$ Hankel matrix is given by

$$H_N = \begin{bmatrix} \hat{G}_1 & \ldots & \hat{G}_N \\ \vdots & & \\ \hat{G}_N & \ldots & \hat{G}_{2N-1} \end{bmatrix}$$

where $N=(r+1)/2$, which is assumed to be an integer. The estimated finite Hankel matrix can be factorized by means of the SVD with $p \leqslant q$ as

$$H_N = U(\Sigma\, 0)\begin{pmatrix} V'_1 \\ V'_2 \end{pmatrix} = U \Sigma V'_1 = \sum_{i=1}^{pN} \sigma_i u_i v'_i$$

where $U=(u_1, \ldots, u_{pN})$ is a $pN \times pN$ orthogonal matrix; $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_{pN})$ with singular values $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_{pN} > 0$ and $V=(V_1\, V_2)=(v_1, \ldots, v_{pN}, v_{pN+1}, \ldots, v_{qN})$ is a $qN \times qN$ orthogonal matrix. Because the Hankel matrix consists of estimates (noisy) impulse response functions $\{\hat{G}_i\}$ it is to be expected that $r(H_N) = pN$, i. e. full row rank. A balanced realization for the "systematic" part

$$\bar{y}_t = \sum_{i=0}^{r} \hat{G}_i u_{t-i} + \hat{C}$$

is given by $\Sigma(\hat{A}_b, \hat{B}_b, \hat{C}_b, \hat{D})$:

$$x_{t+1} = \hat{A}_b x_t + \hat{B}_b u_t$$
$$\bar{y}_t = \hat{C}_b x_t + \hat{D} u_t$$

where dim $(x_t) = p\,N$; $\hat{A}_b = Q^+ Q^\uparrow$; $\hat{B}_b$ = first q columns of P, $\hat{C}_b$ = first p rows of Q and $\hat{D} = \hat{G}_0$. Here $Q^+ = (Q'Q)^{-1} Q'$ is the pseudo-inverse of Q where $Q = U\Sigma^{1/2}$ and $P = \Sigma^{1/2} V_1'$. The shifted matrix

$$Q^\uparrow = \left[\frac{Q_2}{0}\right] \quad \text{with} \quad Q = \left[\frac{Q_1}{Q_2}\right]$$

where Q_1 consists of the first p rows of Q.

Suppose now that the last $(p\,N - k)$ singular values are "close to zero", i.e. $\sigma_{k+i} \approx 0$, $i \geqslant 1$.

A k-th order approximation is then given by $\Sigma(\hat{A}_{11}, \hat{B}_1, \hat{C}_1, \hat{D})$, with

$$\hat{A}_b = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}; \qquad \hat{B}_b = \left(\frac{\hat{B}_1}{\hat{B}_2}\right); \qquad \hat{C}_b = (\hat{C}_1 \,|\, \hat{C}_2)$$

where $\hat{A}_{11}$ is a $k \times k$ matrix; $\hat{B}_1$ a $k \times q$ matrix and $\hat{C}_1$ a $p \times k$ matrix. The dimension of the state vector in the approximated model is k. Including the disturbance term δ_t we have as stochastic k-th order approximated state space model $\Sigma_s(\hat{A}_{11}, \hat{B}_1, \hat{C}_1, \hat{D})$:

$$x_{t+1} = \hat{A}_{11} x_t + \hat{B}_1 u_t$$
$$y_t = \hat{C}_{11} x_t + \hat{D} u_t + \delta_t$$

where $\dim(x_t) = k$. The steady state Kalman filter for this model is $\Sigma_k(\hat{A}_{11}, \hat{B}_1, \hat{C}_1, \hat{D})$:

$$\hat{x}_{t+1\,|\,t} = \hat{A}_{11} \hat{x}_{t\,|\,t-1} + \hat{B}_1 u_t + \hat{A}_{11} K e_t$$

initial condition $\hat{x}_0 = \mu$.

$$y_t = \hat{C}_1 \hat{x}_{t\,|\,t-1} + \hat{D} u_t + e_t$$

where K is the steady state Kalman gain and e_t the prediction error given by

$$e_t = y_t - \hat{y}_{t\,|\,t-1} = y_t - \hat{C}_1 \hat{x}_{t\,|\,t-1} - \hat{D} u_t$$

In section four we consider the estimation of K and μ.

3.2. Sample Models with Stochastic Inputs

When the input $\{u_t\}$ is stochastic (a stationary process) the following procedure is suggested.

Step I: pre-whiten $\{u_t\}$ by fitting an ARMA (p, q) model

$$\hat{A}(L)\, u_t = \hat{B}(L)\, \hat{\eta}_t$$

along the lines of Box and Jenkins [1970]. Here the sequence $\{\hat{\eta}_t\}$ are input prediction errors (realizations of η_t) which is a Gaussian white noise process with

$$\hat{\eta}_t = u_t - \hat{u}_{t \mid t-1}$$

where the prediction

$$\hat{u}_{t \mid t-1} = -\hat{A}_1 u_{t-1} - \ldots - \hat{A}_p u_{t-p} + \hat{B}_1 \hat{\eta}_{t-1} + \ldots + \hat{B}_q \hat{\eta}_{t-q}.$$

the standard white noise input prediction is

$$\tilde{\hat{\eta}}_t = \hat{F}_\eta^{-1} \hat{\eta}_t$$

where the estimated covariance $\hat{V}_\eta$ is decomposed as $\hat{V}_\eta = \hat{F}_\eta \hat{F}'_\eta$.

Step II: Estimate the impulse response function $\psi_1, \psi_2, \ldots$ from the model

$$y_t = \psi(L) \tilde{\hat{\eta}}_t + \delta_t$$

and proceed as before, but now instead of using u_t the input prediction errors $\tilde{\hat{\eta}}_t$.

A steady state Kalman filter is given by

$$\hat{x}_{t+1 \mid t} = \hat{A}_b \hat{x}_{t \mid t-1} + \hat{B}_b \tilde{\hat{\eta}}_t + \hat{A}_b K e_t \qquad \text{initial value } \hat{x}_0 = \mu$$

$$y_t = \hat{C}_b \hat{x}_{t \mid t-1} + \hat{D}_b \tilde{\hat{\eta}}_t + e_t.$$

The only difference with the non-stochastic input case is that instead of using the input sequence $\{u_t\}$ in the filter for the stochastic input case the pre-whitened input (input prediction errors) sequence $\{\tilde{\hat{\eta}}_t\}$ is being used. The steady-state Kalman filter is now a linear function of the input prediction error sequence $\{\tilde{\hat{\eta}}_t\}$ and the output prediction error sequence $\{e_t\}$.

From section 2.3 we have that the mutual information between the future output y_t^+ and past input innovation $\tilde{\eta}_t^-$ for all t is given by

$$I(y_t^+, \tilde{\eta}_t^-) = -\frac{1}{2} \sum_{i=1}^{n} \log(1 - \sigma_i^2)$$

where σ_i, $i = 1, \ldots, n$ are the population canonical correlation coefficients. Let $H_\psi(T)$ be a finite ML-estimate of the Hankel (covariance) matrix H_ψ as given in section 2.2 based on T observations and such that

$$\hat{H}_\psi(N, T) = \begin{bmatrix} \hat{\psi}_1(T) & \ldots & \hat{\psi}_N(T) \\ \vdots & & \vdots \\ \hat{\psi}_N(T) & \ldots & \hat{\psi}_{2N-1}(T) \end{bmatrix}$$

where $\hat{\psi}_i(T)$, $i = 1, \ldots, 2N-1$ with $N < (T+1)/2$ are ML-estimates of ψ_i, where it is assumed that $\text{plim}(\hat{H}_\psi(N, T)) = H_\psi$ and

$r(\mathrm{H}_\psi) = n < \inf_{\mathrm{N,\,T}} r(\hat{\mathrm{H}}_\psi(\mathrm{N,\,T}) \leqslant \mathrm{N}p$ where $p \leqslant q$ is the dimension of the output and input respectively. According to ANDERSON [1958] we have that the canonical correlation coefficients $\hat{\sigma}_i$, $i = 1, \ldots, \mathrm{N}p$ based ou $\hat{\mathrm{H}}_\psi$ (N, T) are ML estimates of σ_i. With the test statistic

$$w := -[\mathrm{T} - \frac{1}{2}(\mathrm{N}p + \mathrm{N}q + 1)] \log \prod_{i=k+1}^{\mathrm{N}p} (1 - \hat{\sigma}_i^2)$$

the null hypothesis $\sigma_{k+1} = \sigma_{k+2} = \ldots = \sigma_{\mathrm{N}p} = 0$ can be tested. Under the null hypothesis w is asymptotically distributed as

$$w \sim \chi^2_{(\mathrm{N}p-k)\,(\mathrm{N}q-k)}$$

see BARTLETT [1941].

3.3. Prediction Error Estimation

In section two steady-state Kalman filters have been considered with parameters (A_b, B_b, C_b, D, V_δ, K, μ) where A_b, B_b, C_b can be calculated using the principal Hankel component procedure and K, μ and V_δ are respectively the steady-state gain, the starting value of the state-vector and the error covariance matrix of δ_t. It was assumed that $\{\mathrm{G}_i\}$ and V_δ were known and that the "true" rank of the Hankel matrix equals n.

Because the Kalman predictor $x_{t\,|\,t-1}$ is optimal in the sense that the error-covariance Σ is minimum we have that the prediction error covariance of the prediction-error $\{e_t\}$, given by

$$\mathrm{V}_e = \mathrm{C}_b \Sigma \mathrm{C}_b' + \mathrm{V}_\delta$$

is minimum.

Let $\theta_0 \in \Theta \subset \mathrm{R}^s$ be the "true" parameter vector in Θ open such that $\mathrm{V}_e(\theta_0)$ is minimum, where θ_0 consists of (A_b, B_b, C_b, D, V_δ, K, μ). Assume that the true dimension of the state-vector is known and that the Kalman predictors

$$y_{t\,|\,t-1}(\theta) = g\,(y^{(t-1)}, u^{(t)}, \theta)$$

and

$$y_{t\,|\,t-1}(\theta_0) = g\,(y^{(t-1)}, u^{(t)}, \theta_0)$$

are such that for $\theta \neq \theta_0$ with $\theta \in \Theta$ with propability one $y_{t\,|\,t-1}(\theta) \neq y_{t\,|\,t-1}(\theta_0)$ for all t. Here $y_{t\,|\,t-1}(\theta)$ denotes a Kalman predictor with parameter vector $\theta \in \Theta$ and where $y^{(t-1)} = (y_{t-1}, y_{t-2}, \ldots)$ and $u^{(t)}$ are the previous observations.

Consider the prediction-error criterion

$$\mathrm{J}_\mathrm{T}(\theta) = \log \det \mathrm{D}_\mathrm{T}(\theta)$$

with as sample prediction-error covariance

$$\mathrm{D}_\mathrm{T}(\theta) = \mathrm{T}^{-1} \sum_{j=1}^{\mathrm{T}} e_j(\theta)\, e_j'(\theta).$$

with prediction-error

$$e_j(\theta) = y_j - y_{j|j-1}(\theta)$$

Non-linear minimization of $J_T(\theta)$ with respect to θ yields a prediction-error (ML) estimate $\hat{\theta}_T$, which is strongly consistent. See for details GOODWIN and PAYNE [1977] and for numerical aspects MEHRA [1974].

4. Estimated Balanced Realizations and Predictions of Three Macroeconometric Equations

In the sequel we consider three equations of a small econometric model, called the Grecon model representing the Dutch economy, see DIETZENBACHER and co-workers [1984].

Private consumption

$$c_t = f_1(L_{B-(1/2)}, NL_{B-(1/2)}, c_c).$$

Private Gross investments

$$i_m = f_2(v'_{-(1/2)}, K_{-1.5}).$$

Price of consumption

$$p_c = f_3(HB, p_{m-(1/2)}, T_k'')$$

where f_1, f_2 and f_3 are linear functions and

c total private consumption;
$L_{B-(1/2)}$ disposable income of private persons from wages and social payments (lagged half a year);
NL_B disposable income of private persons excluding wages and social payments;
c_c consumption credits;
i_m private gross fixed investments, excluding dwellings;
v' total expenditures minus increase of inventories minus exports of services minus non material government consumption (output of commodities);
K gross profit per unit production;
p_c price of consumption;
HB proportion of average gross wages per manyear and the productivity of labour;
p_m price of imports;
T_k'' difference of indirect taxes as a percentage of the output of commodities.

The procedure is the following.

Step I: Estimate the impulse response function $\{G_i\}$ in the equation

$$y_t \approx \sum_{i=0}^{r} G_i u_{t-i} + \text{Const.} + \varepsilon_t$$

where r is the maximum lag given the length of the time series. We have taken $r=5$ and the constant has been set equal to zero.

Step II: Form the finite Hankel with dimensions $Np \times Nq = 3 \times (3 \times q)$ because $N=(r+1)/2$ and where q is the number of inputs.

Step III: Apply a singular value decomposition (SVD) of the finite Hankel.

Step IV: From the SVD a balanced realization can be obtained.

As an example consider the Hankel matrix for investments (sample period 1952-1982).

$$H = \begin{bmatrix} -0.406 & 1.177 & 0.737 & -0.053 & -1.256 & -0.005 \\ 0.737 & -0.053 & -1.256 & -0.005 & 0.482 & -0.046 \\ -1.256 & -0.005 & 0.482 & -0.046 & -0.344 & -0.176 \end{bmatrix} = U \Sigma V'$$

with

$$U = \begin{bmatrix} -0.693 & -0.702 & 0.162 \\ 0.562 & -0.386 & 0.732 \\ -0.451 & 0.598 & 0.662 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2.460 & 0 & 0 \\ 0 & 1.238 & 0 \\ 0 & 0 & 0.642 \end{bmatrix}$$

and

$$V' = \begin{bmatrix} 0.513 & -0.343 & -0.583 & 0.022 & 0.527 & 0.023 \\ -0.607 & -0.654 & 0.207 & 0.009 & 0.397 & -0.068 \\ -0.558 & 0.231 & -0.749 & -0.067 & -0.122 & -0.235 \end{bmatrix}$$

The singular values are $\sigma_1 = 2.460$; $\sigma_2 = 1.237$ and $\sigma_3 = 0.642$. A balanced realization is given by

$$\begin{bmatrix} x1(t+1) \\ x2(t+1) \\ x3(t+1) \end{bmatrix} = \begin{bmatrix} -0.234 & 0.309 & 0.498 \\ -0.364 & -0.715 & -0.292 \\ -0.015 & -0.355 & 0.621 \end{bmatrix} \times \begin{bmatrix} x1(t) \\ x2(t) \\ x3(t) \end{bmatrix} + \begin{bmatrix} 0.199 & -0.755 \\ -0.804 & 0.381 \\ 0.338 & 0.502 \end{bmatrix} \times \begin{bmatrix} u1(t) \\ u2(t) \end{bmatrix}$$

$$\bar{y}(t) = [-1.220 \quad 0.598 \quad -0.310] \begin{bmatrix} x1(t) \\ x2(t) \\ x3(t) \end{bmatrix} + [1.522 \quad -0.571] \begin{bmatrix} u1(t) \\ u2(t) \end{bmatrix}.$$

$$A_b = \begin{bmatrix} -0.234 & 0.309 & 0.498 \\ -0.364 & -0.715 & -0.292 \\ -0.015 & -0.355 & 0.621 \end{bmatrix}, \quad B_b = \begin{bmatrix} 0.199 & -0.755 \\ -0.804 & 0.381 \\ 0.338 & 0.502 \end{bmatrix}$$

$$C_b = [-1.220 \quad 0.598 \quad -0.310] \quad \text{and} \quad D = G_0 = [1.552 \quad -0.571]$$

with respectively $A_b = \Sigma^{-1/2} U' U^{\dagger} \Sigma^{1/2}$ where $U^{\dagger}$ consists of the last two rows of U and the last row consists of zeros; B_b = first $q(=2)$ colums of $\Sigma^{1/2} V'$; C_b = first row of $U \Sigma^{1/2}$ and $D = \hat{G}_0$. Inspection of the singular values may lead to a reduced order model, e. g. if σ_3 is "close to zero" then a reduced order model is with $\dim(x_t) = 2$;

$$\begin{bmatrix} x1(t+1) \\ x2(t+1) \end{bmatrix} = \begin{bmatrix} -0.234 & 0.309 \\ -0.364 & -0.715 \end{bmatrix} \times \begin{bmatrix} x1(t) \\ x2(t) \end{bmatrix} + \begin{bmatrix} 0.199 & -0.755 \\ -0.804 & 0.381 \end{bmatrix} \times \begin{bmatrix} u1(t) \\ u2(t) \end{bmatrix}$$

$$y(t) = (-1.220 \quad 0.598) \begin{bmatrix} x1(t) \\ x2(t) \end{bmatrix} + (1.552 \quad -0.571) \begin{bmatrix} u1(t) \\ u2(t) \end{bmatrix}$$

The following singular values were obtained.

Sample period	Consumption			Investment			Consumption price		
	σ_1	σ_2	σ_3	σ_1	σ_2	σ_3	σ_1	σ_2	σ_3
1952-1979.......	0.526	0.270	0.192	2.633	1.305	0.629	1.390	0.677	0.540
1952-1980.......	0.538	0.301	0.199	2.468	1.235	0.650	1.468	0.735	0.374
1952-1981.......	0.522	0.316	0.199	2.466	1.235	0.650	1.445	0.708	0.380
1952-1982.......	0.471	0.227	0.183	2.460	1.237	0.642	1.411	0.677	0.286

In order to estimate the state vector $\{x_t, t=1, 2, \ldots\}$ the following steady-state Kalman filter has been used

$$x_{t+1\,|\,t} = A_b x_{t\,|\,t-1} + B_b u_t + A_b K e_t \qquad \text{initial condition } x_0 = \mu,$$
$$y_t = C_b x_{t\,|\,t-1} + D u_t + e_t, \qquad \hat{x}_{1\,|\,0} = A_b \mu + B_b u_0.$$

where K is the steady-state gain and $\{e_t\}$ the prediction-error sequence (innovations) which is Gaussian white-noise with $E\{e_t \,|\, \Omega_{t-1}\} = 0$ where $\Omega(t-1) = \{y_{t-1}, \ldots, u_{t-1}, \ldots\}$. Denote by θ the (unknown) parameters of $(A_b, B_b, K, C_b, D, \mu)$ then by minimizing the prediction error criterion

$$V_T(\theta) = \log \det D_T(\theta)$$

where $D_T(\theta) = T^{-1} \sum_{i=1}^{T} w_i^2(\theta)$ is the sample prediction error variance of the sample prediction errors $\{w_i\}$ as realizations of $\{e_i\}$. The prediction error criterion equals in this case minus the log-likelihood. The procedure followed was: take as initial sample period 1952-1979, predict for 1980; reestimate with sample period 1952-1980, predict and so forth. As an example the consumption state space model with $\dim(x_t) = 1$ is given.

The prediction errors were tested against white noise by applying *t*-tests, Durbin-Watson tests, the test of Mehra-Peschon, etc. In all cases there is strong evidence that the prediction error sequences are white noise. Finally *ex post* one-period ahead predictions were calculated using state space models with $\dim(x_t) = 1$.

Consumption

Periode	Model
1952-1979.	$x_{t+1} = -0.731\,x_t + (0.380 \quad -0.298 \quad 0.006)\,u_t + 1.338\,e_t$ $y_t = 0.647\,x_t + (0.541 \quad 0.154 \quad 0.020)\,u_t + e_t$ with $x_0 = 2.134$
1952-1980.	$x_{t+1} = -0.870\,x_t + (0.409 \quad -0.712 \quad 0.006)\,u_t + 1.932\,e_t$ $y_t = 0.337\,x_t + (0.582 \quad 0.154 \quad 0.037)\,u_t + e_t$ with $x_0 = 3.121$
1952-1981.	$x_{t+1} = -0.652\,x_t + (0.660 \quad -0.531 \quad 0.016)\,u_t + 2.270\,e_t$ $y_t = 0.388\,x_t + (0.485 \quad 0.167 \quad 0.037)\,u_t + e_t$ with $x_0 = 4.241$
1952-1982.	$x_{t+1} = -0.696\,x_t + (6.150 \quad -5.059 \quad 0.074)\,u_t + 30.269\,e_t$ $y_t = 0.023\,x_t + (0.561 \quad 0.117 \quad 0.039)\,u_t + e_t$ with $x_0 = 24.843$

Consumption

Year	State-space model	Realized Value	Grecon
1980.	0.127 (0.030)	0.097	1.700 (1.603)
1981.	−0.915 (1.597)	−2.512	1.000 (3.512)
1982.	−1.243 (0.078)	−1.321	−0.800 (0.521)
1983.	−2.073 (1.866)	−0.207	0.700 (0.907)

Investments

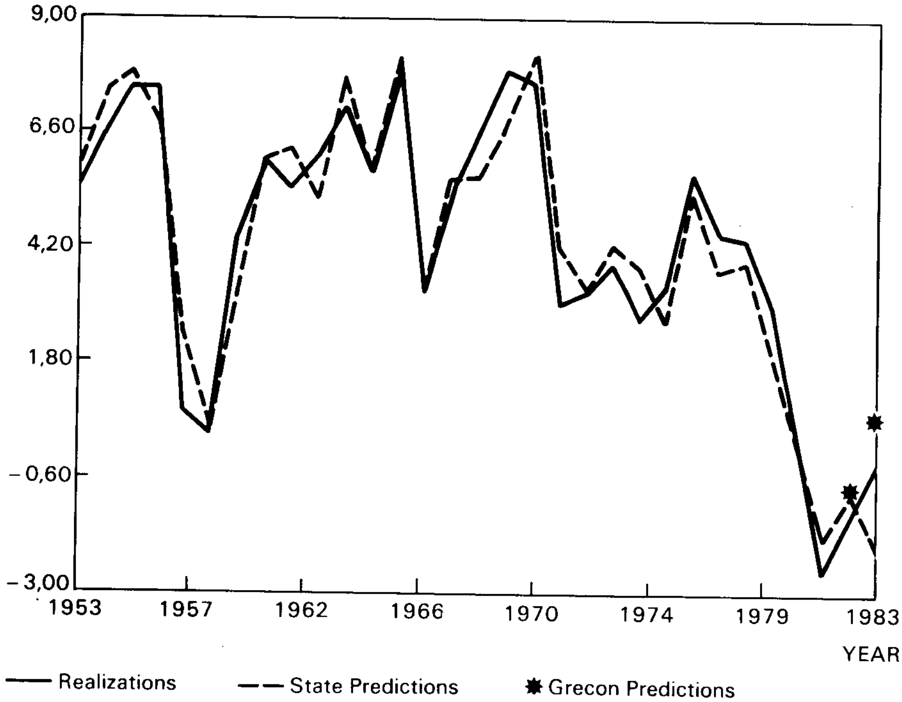
Year	State-space model	Realized Value	Grecon
1980.	−3.169 (1.549)	−4.718	−17.800 (13.082)
1981.	−5.790 (6.883)	−12.673	−28.900 (16.227)
1982.	−2.489 (0.609)	−1.880	1.500 (3.380)
1983.	8.963 (7.676)	1.287	7.300 (6.013)

Consumption price

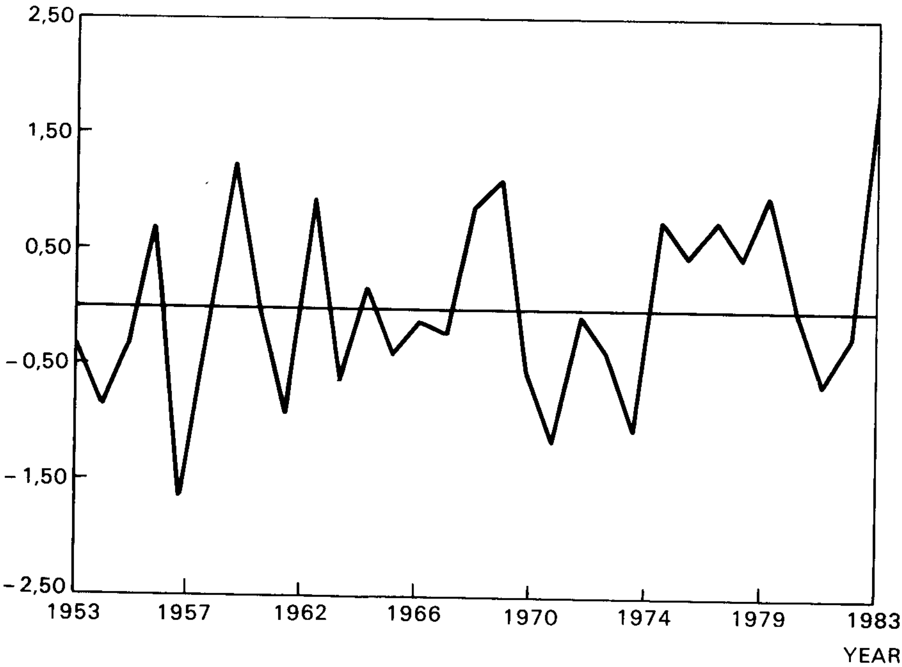
Year	State-space model	Realized Value	Grecon
1980.	5.527 (0.856)	6.383	3.000 (3.383)
1981.	6.399 (0.399)	6.000	3.800 (2.200)
1982.	6.320 (0.660)	5.660	0.100 (5.560)
1983.	2.336 (0.343)	2.679	3.600 (0.921)

The so-called "Grecon" predictions are the *ex post* predictions obtained by using a small econometric model as described in DIETZENBACHER [1984]. Between brackets are the prediction errors in absolute values. It can be seen that the state space models gave in the average better *ex post* predictions. Further research is planned to obtain prediction confidence intervals and prediction-error measures, e. g. Teil's prediction error measure.

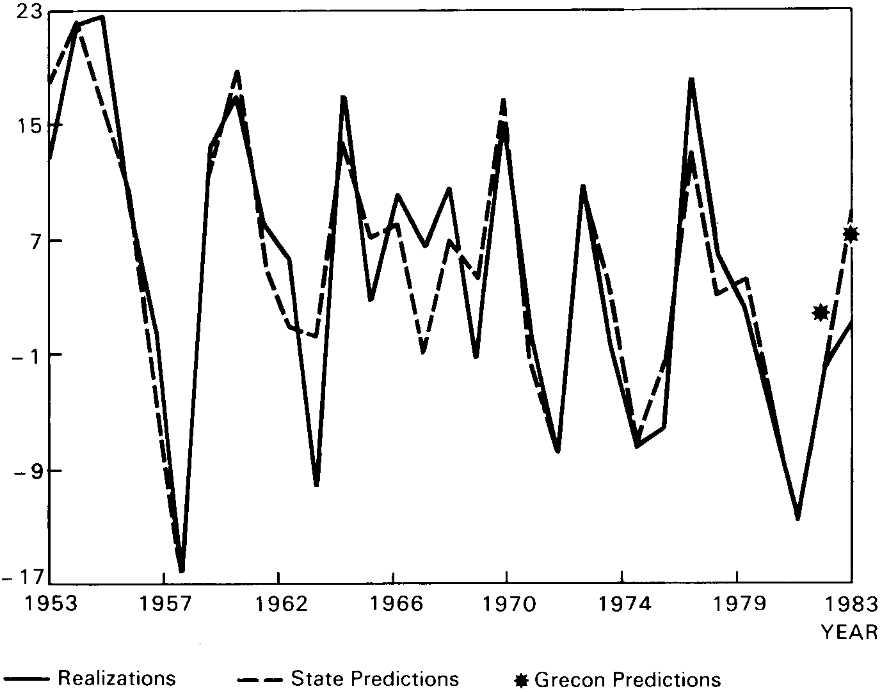
Consumption (Ex Post *Predictions)*



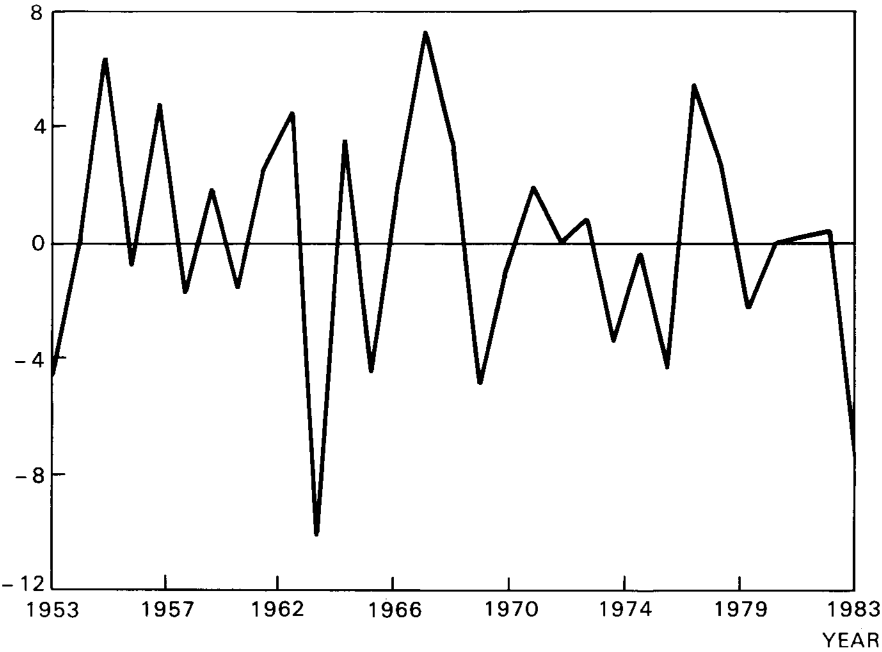
(state) Prediction-Errors



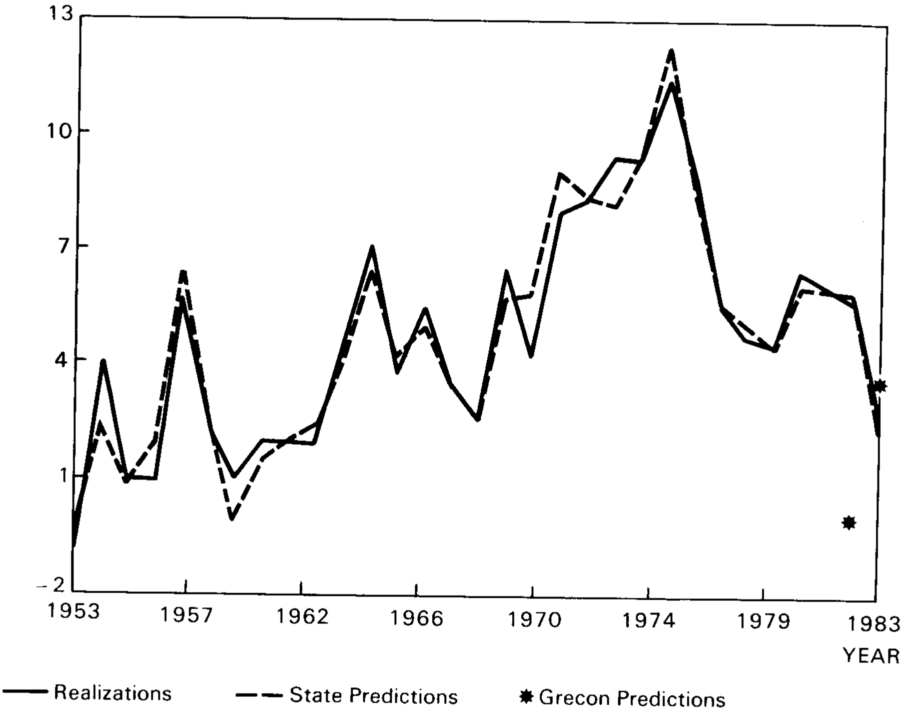
Investments (Ex Post *Predictions)*



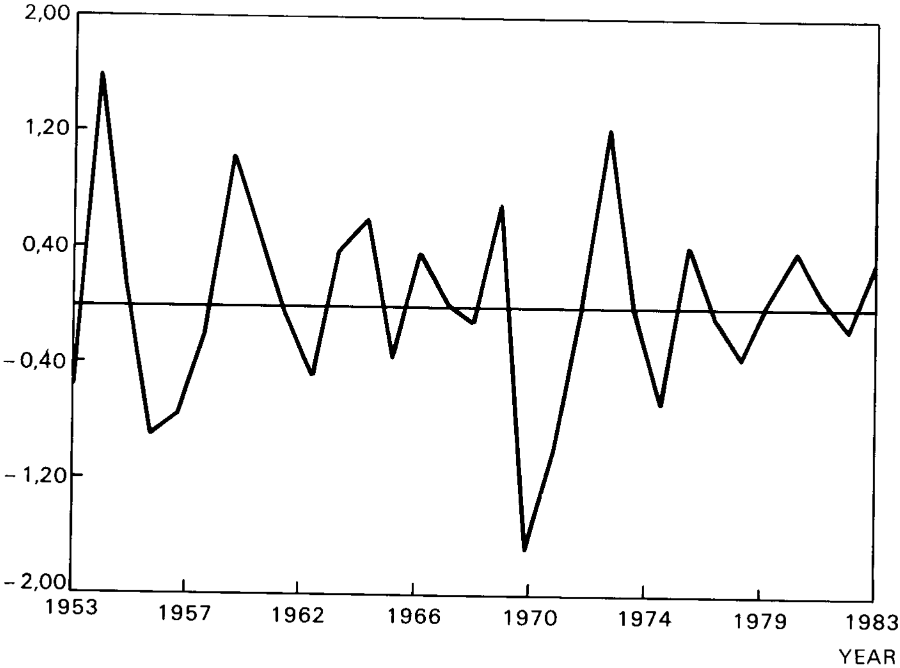
(State) Prediction-Errors



***Consumptionprice* (Ex Post Predictions)**



(State) Prediction-Errors



• References

Anderson, T. W. (1958). — *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York.

Aoki, M. (1983). — Notes on Economic Time Series Analysis: System Theoretic Perspectives", *Lecture Notes in Economics and Mathematical Systems*, Springer Verlag.

Aoki, M. (1986). — "State Space Modelling of Time Series: Examples of US-Japan Interaction of Some Macro Economic Time Series", *Paper presented* at the conference on Modelling Dynamic Systems, Paris.

Aoki, M. (1987). — *State Space Modelling of Time Series*, Springer-Verlag.

Bartlett, M. S. (1941). — The Statistical Significance of Canonical Correlations", *Biometrika.*

Bertsekas, D. P. (1976). — *Dynamic Programming and Stochastic Control*, Academic Press.

Box, G. E. P. and Jenkins G. M. (1970). — *Time Series Analysis Forecasting and Control*, Holden Day.

Brown, R. L., Durbin J. and Evans J. M. (1975). — "Techniques for Testing the Constancy of Regression Relationships over Time, *Journal of Royal Statistical Society*, B37.

Desai, U. B. and Pal, D. (1982). — *A Realization Approach to Stochastic Model Reduction*, Dept. of Electronical Engineering, Washington State University Pullsman, WA 99164.

Dietzenbacher, H. W. A., de Jong V. J., Kooyman M. A., Steerneman A. G. M. and Voorhoeve W. — (1984). — "Het model Grecon 84-D en de voorspellingen voor 1984", *Rapport Econometrisch Instituut.*

Gelfand, I. M. and Yaglom, A. M. (1959). — "Calculation of the Amount of Information about a Random Function Contained in Another Such Function", *American Mathematical Society Transl.*, (2), 12.

Glover, K., "All Optimal Hankel-norm Approximation of Linear Multivariable Systems and their L^{∞}-error Bound", *International Journal of Control*, 39, No. 6, p. 1115-1193.

Goodwin, G. C. and Payne R. L. (1977). — *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, 1977.

Kailath, T. (1980). — *Linear Systems*, Prentice Hall.

Kung, S. Y. and Lin D. W. (1981). — Optimal Hankel-norm Reduction: Multivariable Systems", *IEEE Trans. on Automatic Control*, AC-26, No. 4.

Ljung, L. and Söderström T. (1983). — *Theory and Practice of Recursive Identification*, MIT Press.

Maddala, G. S. (1977). — *Econometrics*, McGraw-Hill.

Mehra, R. V. and J. Peschon (1971). — "An Innovation Approach to Fault Detection and Diagnosis and Dynamic Systems", *Automatica*, 7, 1971.

Mehra, R. K. (1974). — Identification in Control and Econometrics; Similarities and Differences", *Annals of Economic and Social Measurement*, 3/1.

Moore, B. C. (1981). — "Principal Component Analysis in Linear Systems: Controllability, Observability in Model Reduction". *IEEE Trans. on-Automatic Control*, AC-26, No 1.

Otter, P. W. (1978). – "The Discrete Kalman Filter Applied to Linear Regression Models, Statistical Considerations and an Application", *Statistica Neerlandica,* 32, No. 1.

Otter, P. W. (1985). – "Dynamic Feature Space Modelling, Filtering and Self-Tuning Control of Stochastic Systems", *Lecture Notes in Economics and Mathematical Systems,* Springer Verlag, 246.

Otter, P. W. (1986). – "Dynamic Structural Systems under Indirect Observation: Identifiability and Estimation Aspects from a System Theoretic Perspective", *Psychometrika,* 51, No. 3.

Picci, G. (1982). – "Some Numerical Aspects of Multivariable Systems Identification", *Mathematical Programming Study,* 18, North-Holland.

Sage, A. P. and Melsa J. L. (1971). – *Estimation Theory with Applications to Communications and Control,* Mc Graw-Hill.

Silverman, L. M. and Bettayeb, M. (1980). – "Optimal Approximation of Linear Systems", *Proceedings JACC,* San Francisco, CA.

Watson, M. W. and Engle R. F. (1983). – Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC and Varying Coefficient Regression Models, *Journal of Econometrics,* 23, No. 3.

Watson, M. W. and Engle R. F. (1985). – "Applications of Kalman Filtering in Econometrics", *Paper presented* at the World Congress of the Econometric Society, Cambridge Mass.

Willems, J. C. (1984). – "Dynamical Systems", *Lecture Notes,* Mathematical Institute Groningen.