Lecture Notes on the Theory, Calibration & Estimation of Dynamic Stochastic General Equilibrium Models*

Jim Malley
Department of Economics
University of Glasgow
Glasgow G12 8RT
Tel: +44 (0)141 330 4618
Fax: +44 (0)141 330 4940
E-mail: j.malley@socsci.gla.ac.uk

November 30, 2004

Contents

*I would like to thank Konstantinos Angelopoulos, Louis Christofides, Campbell Leith, Apostolis Philippopoulos and Ulrich Woitek for many helpful comments and suggestions. I would also like to thank the Center for Economics Studies, University of Munich for supporting my autumn 2004 sabbatical visit during which these lectures were further developed. Finally I would like to thank the postgraduate students at the Institute for Advanced Studies in Vienna and Cyprus for their lively interaction during the lectures (not to mention the "Heuriger") and for helping to spot the typos. Hopefully these have by now all been removed but if any errors remain, the usual disclaimer applies.

1. Introduction

The purpose of this set of lectures is to set out from first principles the theory, calibration and estimation of a dynamic stochastic general equilibrium model (DSGE). To this end we will follow Ireland's (2004, 2003 and 1999) treatment of Hansen's (1985) real business cycle (RBC) model incorporating an exogenous technology shock. Ireland's work illustrates a new method for estimating this class of model. His approach avoids the problem of singularity associated with the maximum likelihood (ML) estimation of DSGE models with one random shock. To circumvent this difficulty Ireland follows Sargent (1989), Altug (1989), McGratten (1994), Hall (1996) and McGratten, *et al.* (1997) and first adds scalar *AR(1)* errors to the observation equations of the state-space representation of the RBC model. However in contrast to these authors, Ireland allows for cross equation correlation between these errors. This is achieved via a vector *AR(1)* structure for the errors. Hence this hybrid representation allows the estimation not only of the underlying structural parameters dictated by the theory but also of movements and co-movements in the errors which the theory and scalar AR processes cannot explain.

Our discussion of model specification will first cover a description of the economic environment in which the representative agent makes consumption, labour supply and production decisions as well as the predetermined processes describing technological progress and the evolution of the capital stock. Second we will examine the optimization undertaken by the representative consumer/producer to determine its equilibrium consumption and investment allocations. This will require discussion of (i) the equilibrium conditions of the model; (ii) the transformations required to render the system stationary; (iii) the steady-state; and (iv) the log-linearized representation of the equilibrium conditions. This analysis allows us to succinctly describe the behavior of the model's stationary magnitudes as they fluctuate about their steady-state values in response to technology shocks.

Solution of the system of log-linearized dynamic and static equilibrium conditions is required to undertake the ML estimation of the model as set out in Ireland. This is achieved by applying the method developed by Blanchard and Kahn (1980) and yields a state-space representation where expectations for all the model's forward looking equations have been "solved out". Accordingly the model's state and control variables depend on the underlying structural parameters, the predetermined state variables and exogenous shocks. To facilitate a deeper understanding of the model's structure and solution we will calibrate it using parameters from Ireland (1999) prior to estimating the model from 1948Q1 to 2002Q2.

The estimation exercise will make use of Ireland's MATLAB code and data to replicate the results in his (2004) *JEDC* paper. Estimation of the state space representation will be carried out using the Kalman filter to maximize the unconstrained and constrained likelihood functions of the RBC model. In addition, the estimated model will be empirically evaluated by examining (i) impulse responses; (ii) variance decompositions; (iii) stability tests; and (iv) measures of forecast accuracy.

Detailed appendices have been added to review the basic tools required to understand and manipulate the economic and econometric models. These include: (i) alternative methods of log-linearizing non-linear models; (ii) alternative methods of solving first-order linear difference equations; (iii) a method to solve systems of linear difference equations; (v) applying the Kalman filter to maximize the (unconstrained and constrained) likelihood function of a state-space model; (vi) a short description of the MATLAB code by Peter Ireland which can be used to estimate, solve, and evaluate his hybrid version of Hansen's RBC model.

2. Economic Environment

2.1. Consumption and labour supply

A representative consumer's expected lifetime utility is given by

$$E\sum_{t=0}^{\infty}\beta^t[\ln(C_t)-\gamma H_t] \tag{2.1}$$

where C_t is consumption, H_t is hours worked and $0<\beta<1$ and $\gamma>0$ are the subjective rate of time preference and the parameter governing the linearity of utility in hours respectively. The latter is motivated in Hansen (1985) and Rogerson (1988) by assuming that the economy is populated by many individual consumers each of whom either work full time or remain unemployed. In other words, considerations pertaining to part-time and over-time working are ignored.

2.2. Production

The representative consumer/producer produces output according to the following constant returns to scale Cobb-Douglas production function which incorporates both Hick's neutral and labour augmenting technological progress

$$Y_t = A_tK_t^{\theta}(\eta^tH_t)^{1-\theta} \tag{2.2}$$

where Y_t is output, A_t is the term representing neutral technological progress, K_t is the capital stock, $\eta > 1$ is the gross rate of labor-augmenting technological progress and $0 < \theta < 1$ is the elasticity of output with respect to capital.

2.3. Technological progress

Technology shocks are assumed to follow an exponential first-order autoregressive process,

$$\begin{aligned} A_t &= A^{(1-\rho)} A_{t-1}^{\rho} e^{\varepsilon_t} \\ \ln(A_t) &= (1-\rho)\ln(A) + \rho \ln(A_{t-1}) + \varepsilon_t \end{aligned} \tag{2.3}$$

where $A > 0$ is a constant, $-1 < \rho < 1$ is the first-order autoregressive parameter, and ε_t is the stochastic error term $\varepsilon_t \sim N(0, \sigma^2)$, $cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$.

2.4. Aggregate resource constraint

In each period $t = 0, 1, 2, ...$ the representative consumer/producer consumes and invests its output according the following aggregate consistency condition

$$Y_t = C_t + I_t \tag{2.4}$$

where I_t is investment.

2.5. Law of motion for capital accumulation

According to (2.1) current consumption yields instantaneous utility whereas current investment provides future utility via its affect on capital accumulation, e.g.

$$K_{t+1} = (1-\delta)K_t + I_t \tag{2.5}$$

and hence future output and consumption via (2.2) and 2.4), where $0 < \delta < 1$ is the depreciation rate of capital.

2.6. Variables and parameters

The model set out above has six endogenous variables (two *predetermined* state variables, i.e. K_t and A_t plus four *non-predetermined* control variables, i.e. Y_t, C_t, I_t, H_t), one exogenous variable, i.e. ε_t and eight parameters. All the parameters *except* ρ and σ^2 are known as deep or structural parameter since they have a precise economic interpretation. The parameters ρ and σ^2 in contrast are specific to the predetermined exogenous statistical process assumed to drive technology. We will see below (Section 3.4) that ρ and σ^2 have no impact on the model's steady-state.

Variables

Y_t	output
C_t	consumption
I_t	investment
H_t	hour worked
K_t	capital stock
A_t	Hick's neutral technological progress
ε_t	innovation to technology

Parameters

$0 < \beta < 1$	subjective rate of time preference
$\gamma > 0$	parameter governing linearity of utility in hours
$\eta > 1$	gross rate of labour augmenting technological progress
$0 < \theta < 1$	elasticity of Y_t with respect to the K_t
$A > 0$	constant term in the process describing A_t
$0 < \delta < 1$	depreciation rate of K_t
$-1 < \rho < 1$	1st-order autoregressive parameter in the A_t process
$\sigma^2 > 0$	error variance in the A_t process

3. Characterization of Equilibrium Allocations

3.1. Optimization

In this setup, the representative consumer chooses $\{Y_t, C_t, I_t, H_t, K_{t+1}\}_{t=0}^{\infty}$ to maximize utility subject to the production function (2.2), the law of motion for capital (2.5) and the aggregate resource constraint (2.4), for all $t = 0, 1, 2, ...$ We can simplify this problem by substituting constraints (2.2) and (2.4) into (2.5) to obtain

$$K_{t+1} = (1-\delta)K_t + A_t K_t^{\theta}(\eta^t H_t)^{1-\theta} - C_t. \tag{3.1}$$

Given these substitutions the household can now choose $\{C_t, H_t, K_{t+1}\}_{t=0}^{\infty}$ to maximize utility (2.1) subject to the consolidated constraint (3.1). Accordingly the Lagrangian can be written as follows,

$$\Lambda = E\sum_{t=0}^{\infty}\left\{\beta^t[\ln(C_t) - \gamma H_t] + \beta^t\lambda_t[(1-\delta)K_t + A_t K_t^{\theta}(\eta^t H_t)^{1-\theta} - C_t - K_{t+1}]\right\}$$

To appreciate the nature of the optimization in which the representative consumer chooses the optimal time paths for $\{C_t,\ H_t,\ K_{t+1}\}_{t=0}^{\infty}$ we can unpack the infinite sums in above Lagrangian so that time periods t and $t+1$ appear explicitly, e.g.

$$\begin{aligned}\Lambda \;=\; & E\sum_{j=0}^{t-1}\{\beta^j[\ln(C_j)-\gamma H_j]+\beta^j\lambda_j[(1-\delta)K_j+A_jK_j^{\theta}(\eta^jH_j)^{1-\theta}-C_j-K_{j+1}]\}\\ & +E\beta^t[\ln(C_t)-\gamma H_t]+E\beta^t\lambda_t[(1-\delta)K_t+A_tK_t^{\theta}(\eta^tH_t)^{1-\theta}-C_t-K_{t+1}]\\ & +E\beta^{t+1}[\ln(C_{t+1})-\gamma H_{t+1}]+E\beta^{t+1}\lambda_{t+1}[(1-\delta)K_{t+1}+A_{t+1}K_{t+1}^{\theta}(\eta^{t+1}H_{t+1})^{1-\theta}\\ & -C_{t+1}-K_{t+2}]+E\sum_{j=t+2}^{\infty}\{\beta^j[\ln(C_j)-\gamma H_j]+\beta^j\lambda_j[(1-\delta)K_j+A_jK_j^{\theta}(\eta^jH_j)^{1-\theta}\\ & -C_j-K_{j+1}]\}.\end{aligned}$$

The first-order conditions are derived by setting the partial derivatives of Λ with respect to $C_t,\ H_t,\ K_{t+1}$ to zero, i.e.

$$\Lambda_{C_t}=\frac{\beta^t}{C_t}-\beta^t\lambda_t=0 \tag{3.2}$$

$$\Lambda_{H_t}=-\beta^t\gamma+\beta^t\lambda_t\left[A_tK_t^{\theta}\frac{(1-\theta)(\eta^tH_t)^{1-\theta}}{H_t}\right]=0 \tag{3.3}$$

$$\Lambda_{K_{t+1}}=-\beta^t\lambda_t+E_t\beta^{t+1}\lambda_{t+1}\left[(1-\delta)+A_{t+1}\theta K_{t+1}^{\theta-1}(\eta^{t+1}H_{t+1})^{1-\theta}\right]=0. \tag{3.4}$$

Equation (3.2) implies $\lambda_t=\frac{1}{C_t}$. Substituting λ_t into (3.3) gives

$$0=-\beta^t\gamma+\beta^t\frac{1}{C_t}\left[A_tK_t^{\theta}\frac{(1-\theta)(\eta^tH_t)^{1-\theta}}{H_t}\right]$$

and then making use of (2.2): $Y_t=A_tK_t^{\theta}(\eta^tH_t)^{1-\theta}$ in the previous expression yields

$$\begin{aligned}0 &= -\gamma+\frac{(1-\theta)}{C_t}\left[\frac{Y_t}{H_t}\right]\\ \gamma C_tH_t &= (1-\theta)Y_t.\end{aligned} \tag{3.5}$$

Equation (3.4) and the production function in $t+1$ imply

$$\lambda_t=\beta E_t\lambda_{t+1}\left[(1-\delta)+\theta\frac{Y_{t+1}}{K_{t+1}}\right]. \tag{3.6}$$

From (3.2) we know $\lambda_t=\frac{1}{C_t}$ which implies $\lambda_{t+1}=\frac{1}{C_{t+1}}$. Substituting for λ_t and λ_{t+1} in (3.6) gives

$$1/C_t=\beta E_t\{(1/C_{t+1})[\theta\,(Y_{t+1}/K_{t+1})+1-\delta]\}. \tag{3.7}$$

Therefore our two first-order conditions are given by (3.5) and (3.7).

3.2. Equilibrium conditions

The equilibrium behavior in the model is determined by the production function (2.2), the process describing the evolution of technology (2.3), the aggregate consistency condition (2.4), the capital accumulation relation (2.5) and the two first-order conditions, (3.5) and (3.7), i.e.

$$Y_t = A_t K_t^{\theta}(\eta^t H_t)^{1-\theta} \tag{2.2'}$$

$$\ln(A_t) = (1-\rho)\ln(A) + \rho\ln(A_{t-1}) + \varepsilon_t \tag{2.3'}$$

$$Y_t = C_t + I_t \tag{2.4'}$$

$$K_{t+1} = (1-\delta)K_t + I_t \tag{2.5'}$$

$$\gamma C_t H_t = (1-\theta)Y_t \tag{3.5'}$$

and

$$1/C_t = \beta E_t\{(1/C_{t+1})[\theta\,(Y_{t+1}/K_{t+1}) + 1 - \delta]\} \tag{3.7'}$$

for all $t = 0, 1.2, ...$ Since this model is nonlinear and hence does not have a closed-form solution, an approximate one can be obtained by (i) calculating the stationary representation of the model; (ii) defining the steady-state as one where the stationary variables in the model are constant; (iii) log-linearizing the stationary equilibrium conditions around the steady-state and (iv) applying the methods developed by Blanchard and Kahn (*op cit.*) to solve the log-linearized system of stochastic difference equations.

3.3. Transformed (stationary system)

The equilibrium conditions in Section 3.2 can be transformed so that the system is re-expressed in term of six stationary variables, e.g.

$$y_t = a_t k_t^{\theta} {h_t}^{1-\theta} \tag{2.2''}$$

$$\ln(a_t) = (1-\rho)\ln(A) + \rho\ln(a_{t-1}) + \varepsilon_t \tag{2.3''}$$

$$y_t = c_t + i_t \tag{2.4''}$$

$$\eta k_{t+1} = (1-\delta)k_t + i_t \tag{2.5''}$$

$$\gamma c_t h_t = (1-\theta)y_t \tag{3.5''}$$

and

$$\eta/c_t = \beta E_t\{(1/c_{t+1})[\theta\,(y_{t+1}/k_{t+1}) + 1 - \delta]\} \tag{3.7''}$$

where $y_t = Y_t/\eta^t$, $c_t = C_t/\eta^t$, $i_t = I_t/\eta^t$, $k_t = K_t/\eta^t$, $h_t \equiv H_t$, $a_t \equiv A_t$ and $\eta k_{t+1} = \eta \frac{K_{t+1}}{\eta^{t+1}} = \frac{K_{t+1}}{\eta^t}$.

3.4. Steady-state

In the absence of technology shocks, i.e. $\varepsilon_t = 0$ for all $t = 0, 1, 2...$, the economy converges to a steady-state in which each of the six stationary variables is constant, i.e. $y_t = y$, $c_t = c$, $i_t = i$, $h_t = h$, $k_t = k$, and $a_t = a$.

To solve for a, we can first take the exponential of equation (2.3"): $\ln(a_t) = (1-\rho)\ln(A) + \rho\ln(a_{t-1}) + \varepsilon_t$ at the steady-state

$$\begin{aligned} e^{\ln(a)} &= e^{(1-\rho)\ln(A)+\rho\ln(a)} \\ a &= A^{1-\rho}a^{\rho} \\ a^{1-\rho} &= A^{1-\rho} \\ a &\equiv A. \end{aligned} \tag{3.8}$$

Assuming that the steady-state value for y is known we can use (3.7"): $\eta/c_t = \beta E_t\{(1/c_{t+1})[\theta\,(y_{t+1}/k_{t+1}) + 1 - \delta]\}$ to solve for steady-state k,

$$\begin{aligned} \eta/c &= \beta\left[\frac{1}{c}\frac{\theta y}{k} + \frac{1}{c}(1-\delta)\right] \\ \eta/\beta &= \frac{\theta y}{k} + 1 - \delta \\ k &= \left(\frac{\theta}{\eta/\beta - 1 + \delta}\right) y. \end{aligned} \tag{3.9}$$

Equation (2.5"): $\eta k_{t+1} = (1-\delta)k_t + i_t$ and (3.9) can be used to solve for steady-state i,

$$\begin{aligned} \eta k &= (1-\delta)k + i \\ i &= \eta\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right) y - (1-\delta)\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right) y \\ i &= \left(\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right) y. \end{aligned} \tag{3.10}$$

Equation (2.4"): $y_t = c_t + i_t$ and (3.10) can be used to solve for steady-state c,

$$\begin{aligned} y &= c + i \\ c &= \left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\} y. \end{aligned} \tag{3.11}$$

Equation (3.5"): $\gamma c_t h_t = (1-\theta) y_t$ and (3.11) can be used to solve for steady-state h,

$$\begin{aligned} \gamma c h &= (1-\theta) y \\ h &= \left(\frac{1-\theta}{\gamma}\right)\left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\}^{-1}. \end{aligned} \tag{3.12}$$

Finally we can substitute (3.9) and (3.12) into (2.2"): $y = ak^{\theta}h^{1-\theta}$ which gives

$$
\begin{aligned}
y &= a\left[\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right)y\right]^{\theta}\left[\left(\frac{1-\theta}{\gamma}\right)\left(1 - \frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right)^{-1}\right]^{1-\theta} \qquad (3.13)\\
y^{1-\theta} &= a\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right)^{\theta}\left[\left(\frac{1-\theta}{\gamma}\right)\left(1 - \frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right)^{-1}\right]^{1-\theta}\\
y &= a^{1/(1-\theta)}\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right)^{\theta/(1-\theta)}\left(\frac{1-\theta}{\gamma}\right)\left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\}^{-1}.
\end{aligned}
$$

The above system of equations shows how the steady-state value of our six variables y, c, i, h, k and a depend on the structural or "deep" parameters β, γ, θ, η, δ, and A. As pointed out in Section 2.6 the parameters governing the statistical process driving technology shocks ρ and σ^2 have no effect on the steady-state.

3.5. Linearized system

The non-linear stationary equilibrium conditions in Section 3.3 can next be log-linearized to describe the behavior of the stationary variables as the fluctuate about their constant steady-state values in response to random shocks to technology. A review of alternative methods of taking log-linear approximations is provided in the Appendices (see Review 1). Here we will first take natural logs of all the equations in Section 3.3 and then evaluate the total differential of the resulting equations at the steady-state values defined in Section 3.4.

$$
\begin{aligned}
y_t &= a_t k_t^{\theta} {h_t}^{1-\theta}\\
\ln(y_t) &= \ln(a_t) + \theta\ln(k_t) + (1-\theta)\ln(h_t)\\
\frac{1}{y}dy_t &= \frac{1}{a}da_t + \frac{\theta}{k}dk_t + \frac{(1-\theta)}{h}dh_t \qquad (2.2\text{"'})\\
\widehat{y}_t &= \widehat{a}_t + \theta\widehat{k}_t + (1-\theta)\widehat{h}_t
\end{aligned}
$$

$$
\begin{aligned}
a_t &= A^{(1-\rho)}a_{t-1}^{\rho}e^{\varepsilon_t}\\
\ln(a_t) &= (1-\rho)\ln(A) + \rho\ln(a_{t-1}) + \varepsilon_t\\
\frac{1}{a}da_t &= \rho\frac{1}{a}da_{t-1} + d\varepsilon_t \qquad (2.3\text{"'})\\
\widehat{a}_t &= \rho\widehat{a}_{t-1} + \varepsilon_t, \text{ where}\\
d\varepsilon_t &\simeq \ln(e^{\varepsilon_t}/e^{\varepsilon}) = \varepsilon_t - \varepsilon = \varepsilon_t
\end{aligned}
$$

$$\begin{aligned}
y_t &= c_t + i_t \\
\ln(y_t) &= \ln(c_t + i_t) \\
\frac{1}{y} dy_t &= \frac{1}{c+i}(dc_t + di_t) \\
\widehat{y}_t &= \frac{1}{c+i}\left(\frac{dc_t}{c}c + \frac{di_t}{i}i\right) \\
\widehat{y}_t &= \frac{c}{y}\widehat{c}_t + \frac{i}{y}\widehat{i}_t \qquad (2.4\text{'''}) \\
y\widehat{y}_t &= c\widehat{c}_t + i\widehat{i}_t \\
y\widehat{y}_t &= \left(1 - \frac{\theta(\eta-1+\delta)}{\eta/\beta-1+\delta}\right)y\widehat{c}_t + \left(\frac{\theta(\eta-1+\delta)}{\eta/\beta-1+\delta}\right)y\widehat{i}_t \\
(\eta/\beta-1+\delta)\widehat{y}_t &= [(\eta/\beta-1+\delta) - \theta(\eta-1+\delta)]\,\widehat{c}_t + \theta(\eta-1+\delta)\widehat{i}_t
\end{aligned}$$

$$\begin{aligned}
\eta k_{t+1} &= (1-\delta)k_t + i_t \\
\ln(\eta) + \ln(k_{t+1}) &= \ln[(1-\delta)k_t + i_t] \\
\frac{1}{k}dk_{t+1} &= \frac{1}{(1-\delta)k+i}[(1-\delta)dk_t + di_t] \\
\widehat{k}_{t+1} &= \frac{1}{\eta k}\left[(1-\delta)\frac{dk_t}{k}k + \frac{di_t}{i}i\right] \\
\widehat{k}_{t+1} &= \frac{k}{\eta k}(1-\delta)\widehat{k}_t + \frac{i}{\eta k}\widehat{i}_t \qquad (2.5\text{'''}) \\
\eta\widehat{k}_{t+1} &= (1-\delta)\widehat{k}_t + \frac{\left(\frac{\theta(\eta-1+\delta)}{\eta/\beta-1+\delta}\right)y}{\left(\frac{\theta}{\eta/\beta-1+\delta}\right)y}\widehat{i}_t \\
\eta\widehat{k}_{t+1} &= (1-\delta)\widehat{k}_t + (\eta-1+\delta)\widehat{i}
\end{aligned}$$

$$\begin{aligned}
\gamma c_t h_t &= (1-\theta)y_t \\
\ln(\gamma) + \ln(c_t) + \ln(h_t) &= \ln[(1-\theta)] + \ln(y_t) \\
\frac{1}{c}dc_t + \frac{1}{h}dh_t &= \frac{1}{y}dy_t \qquad (3.5\text{'''}) \\
\widehat{c}_t + \widehat{h}_t &= \widehat{y}_t
\end{aligned}$$

and

$$
\begin{aligned}
\eta/c_t &= \beta E_t\{(1/c_{t+1})[\theta\,(y_{t+1}/k_{t+1}) + 1 - \delta]\} \\
\ln\eta - \ln c_t &= \ln\beta + \ln E_t\{(1/c_{t+1})[\theta(y_{t+1}/k_{t+1}) + 1 - \delta]\} \\
-\frac{1}{c}dc_t &= \frac{1}{(\frac{1}{c})[\frac{\theta y}{k} + 1 - \delta]} E_t\left\{\left[\frac{\theta y}{k} + 1 - \delta\right]\left[-\frac{1}{c^2}dc_{t+1}\right] + \frac{\theta}{c}\left[\frac{1}{k}dy_{t+1} - \frac{y}{k^2}dk_{t+1}\right]\right\} \\
-\hat{c}_t &= E_t\left\{\frac{\frac{\theta y}{k} + 1 - \delta}{(\frac{1}{c}\;)[\frac{\theta y}{k} + 1 - \delta]}\left[-\frac{1}{c^2}dc_{t+1}\right]\right\} + \\
&\quad E_t\left\{\frac{\theta}{c}\frac{1}{(\frac{1}{c}\;)[\frac{\theta y}{k} + 1 - \delta]}\left[\frac{1}{k}dy_{t+1} - \frac{y}{k^2}dk_{t+1}\right]\right\} \\
-\hat{c}_t &= -E_t(\hat{c}_{t+1}) + E_t\left\{\frac{\theta}{[\frac{\theta y}{k} + 1 - \delta]}\frac{1}{k}dy_{t+1}\right\} - E_t\left\{\frac{\theta}{[\frac{\theta y}{k} + 1 - \delta]}\frac{y}{k^2}dk_{t+1}\right\} \\
-\hat{c}_t &= -E_t(\hat{c}_{t+1}) + E_t\left\{\frac{\theta}{[\theta\frac{\frac{\eta}{\beta} - 1 + \delta}{\theta} + 1 - \delta]}\frac{1}{k}dy_{t+1}\right\} - \qquad (3.7\text{"'}) \\
&\quad E_t\left\{\frac{\theta}{[\theta\frac{\frac{\eta}{\beta} - 1 + \delta}{\theta} + 1 - \delta]}\frac{y}{k^2}dk_{t+1}\right\}, \text{ where from (3.9) } \frac{(\eta/\beta - 1 + \delta)}{\theta} = \frac{y}{k} \\
\hat{c}_t &= -E_t(\hat{c}_{t+1}) + E_t\left\{\frac{\theta}{\frac{\eta}{\beta}}\frac{1}{k}dy_{t+1}\right\} - E_t\left\{\frac{\theta}{\frac{\eta}{\beta}}\frac{y}{k^2}dk_{t+1}\right\} \\
\frac{\eta}{\beta}\hat{c}_t &= -\frac{\eta}{\beta}E_t(\hat{c}_{t+1}) + E_t\left\{\theta\frac{1}{k}dy_{t+1}\right\} - E_t\left\{\theta\frac{y}{k^2}dk_{t+1}\right\} \\
\frac{\eta}{\beta}\hat{c}_t &= -\frac{\eta}{\beta}E_t(\hat{c}_{t+1}) + E_t\left\{\theta\frac{y}{k}\frac{dy_{t+1}}{y}\right\} - E_t\left\{\theta\frac{yk}{k^2}\frac{dk_{t+1}}{k}\right\} \\
\frac{\eta}{\beta}\hat{c}_t &= -\frac{\eta}{\beta}E_t(\hat{c}_{t+1}) + \theta\frac{y}{k}E_t\hat{y}_{t+1} - \theta\frac{y}{k}E_t\widehat{k}_{t+1} \\
-\frac{\eta}{\beta}\hat{c}_t &= -\frac{\eta}{\beta}E_t(\hat{c}_{t+1}) + \left(\frac{\eta}{\beta} - 1 + \delta\right)E_t\hat{y}_{t+1} - \left(\frac{\eta}{\beta} - 1 + \delta\right)\hat{k}_{t+1}
\end{aligned}
$$

Given that the derivation for (3.7"') was rather long-winded, let's try the other method set out the Appendices (see Review 1: Taylor approximation). Again consider the stationary first order condition (3.7")

$$\eta/c_t = \beta E_t\{(1/c_{t+1})[\theta(y_{t+1}/k_{t+1}) + 1 - \delta]\}.$$

This equation can be expressed equivalently as

$$\begin{aligned}
\eta/(ce^{\widehat{c}_t}) &= \beta E_t\left\{(1/ce^{\widehat{c}_{t+1}})\left[\theta\left(ye^{\widehat{y}_{t+1}}/ke^{\widehat{k}_{t+1}}\right)+1-\delta\right]\right\} \\
\frac{\eta}{c}e^{-\widehat{c}_t} &= \beta E_t\left\{\left(\frac{1}{c}e^{-\widehat{c}_{t+1}}\right)\left[\theta\left(\frac{y}{k}e^{\widehat{y}_{t+1}-\widehat{k}_{t+1}}\right)+1-\delta\right]\right\}.
\end{aligned}$$

Taking the first order Taylor series expansion of the last expression gives

$$\frac{\eta}{c}(1-\widehat{c}_t)=\beta E_t\left\{\frac{1}{c}(1-\widehat{c}_{t+1})\left[\theta\left(\frac{y}{k}(1+\widehat{y}_{t+1}-\widehat{k}_{t+1})\right)+1-\delta\right]\right\}.$$

From 3.9 we know that $y/k=(\eta/\beta-1+\delta)/\theta$, therefore the previous equation can be rewritten as

$$\begin{aligned}
\eta(1-\widehat{c}_t) &= \beta E_t\{(1-\widehat{c}_{t+1})[\eta/\beta+(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})]\} \\
\frac{\eta}{\beta}(1-\widehat{c}_t) &= E_t\{\frac{\eta}{\beta}(1-\widehat{c}_{t+1})+(1-\widehat{c}_{t+1})[(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})]\} \\
-\frac{\eta}{\beta}\widehat{c}_t &= E_t\{-\frac{\eta}{\beta}\widehat{c}_{t+1}+(1-\widehat{c}_{t+1})[(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})]\} \\
&= E_t\{-\frac{\eta}{\beta}\widehat{c}_{t+1}+(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})-\widehat{c}_{t+1}[(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})]\} \\
&= E_t\{-\frac{\eta}{\beta}\widehat{c}_{t+1}+(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})\} \\
-\frac{\eta}{\beta}\widehat{c}_t &= -\frac{\eta}{\beta}E_t\widehat{c}_{t+1}+\left(\frac{\eta}{\beta}-1+\delta\right)E_t\widehat{y}_{t+1}-\left(\frac{\eta}{\beta}-1+\delta\right)\widehat{k}_{t+1}
\end{aligned}$$

The second to last line in the above result follows since all the cross product terms resulting from multiplying out the expression $-\widehat{c}_{t+1}[(\eta/\beta-1+\delta)(\widehat{y}_{t+1}-\widehat{k}_{t+1})]$ are equal to zero when applying a first-order Taylor series approximation. In other words, we have imposed the certainty equivalence property on the model.

4. Blanchard-Kahn Procedure

4.1. Blanchard-Kahn setup

Equations (2.2"'-2.5"', 3.5"' and 3.7"') now form a system of linear stochastic difference equations which we next need to solve. Let's start by defining a vector $\mathbf{s}_t^0$ which contains the *predetermined* and *non-predetermined* variables of a structural model

$$\mathbf{s}_t^0=\begin{bmatrix}\mathbf{s}_t\\ \mathbf{f}_t\end{bmatrix}$$

where $\mathbf{s}$ is an $(n \times 1)$ vector of *predetermined* and $\mathbf{f}$ is an $(m \times 1)$ vector of *non-predetermined* variables at time t.

Next suppose that $\mathbf{s}_t^0$ depends linearly on $\mathbf{s}_{t-1}^0$ and a vector of exogenous variables, $\mathbf{z}$. We can then describe the structural model in $t+1$ assuming rational expectations as follows

$$E_t \begin{bmatrix} \mathbf{s}_{t+1} \\ \mathbf{f}_{t+1} \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{s}_t \\ \mathbf{f}_t \end{bmatrix} + \mathbf{L}\mathbf{z}_t$$

where $\mathbf{K}$ and $\mathbf{L}$ are $(n+m) \times (n+m)$ and $(n+m) \times k$ coefficient matrices respectively and $\mathbf{z}$ is a $(k \times 1)$ vector of exogenous variables. In this context, Blanchard and Kahn point out that the distinction between *predetermined* and *non-predetermined* variables is critical. A *predetermined* variable $\mathbf{s}_{t+1}$ is a function only of variables known at time t, i.e. of variables in $\mathbf{\Omega}_t$ (the information set at t), so that $\mathbf{s}_{t+1} = E_t(\mathbf{s}_{t+1} \mid \mathbf{\Omega}_t)$ whatever the realization of the variables in $\mathbf{\Omega}_{t+1}$. A *non-predetermined* variable $\mathbf{f}_{t+1}$ can be a function of any variable in $\mathbf{\Omega}_{t+1}$, so that $\mathbf{f}_{t+1} = E_t(\mathbf{f}_{t+1} \mid \mathbf{\Omega}_t)$ only if the realizations of all variables in $\mathbf{\Omega}_{t+1}$ are equal to their expectations conditional on $\mathbf{\Omega}_t$. In other words this distinction defines rational expectations and excludes the possibility that agents know the values of endogenous variables but not the values of the exogenous variables.

To solve the structural model, Blanchard and Kahn first transform the system into Jordan canonical form (see Appendices, Review 3). In other words, they diagonalize the $\mathbf{K}$ matrix such that $\mathbf{K} = \mathbf{M}^{-1}\mathbf{N}\mathbf{M}$, where the diagonal elements of $\mathbf{N}$, which are the roots of $\mathbf{K}$ are ordered by increasing absolute value and the columns of $\mathbf{M}^{-1}$ are the eigenvectors of $\mathbf{K}$. More detail on how these matrices are further decomposed is provided below and in Appendix Review 3. Given this setup Blanchard and Kahn prove the following three propositions.

PROPOSITION 1: *If the number of eigenvalues of* $\mathbf{K}$ *outside the unit circle is equal to the number of non-predetermined variables, then there is a unique solution to the system*[1].

PROPOSITION 2: *If the number of eigenvalues of* $\mathbf{K}$ *outside the unit circle is greater than the number of non-predetermined variables, then there is no solution to the system*[2].

PROPOSITION 3: *If the number of eigenvalues of* $\mathbf{K}$ *outside the unit circle is less*

[1] You will also find this case referred to as "determinate" or "stable" in the literature.

[2] You will also find this case referred to as "unstable" in the literature.

than the number of non-predetermined variables, then there is an infinity of solutions[3].

4.2. Linearized RBC model in matrix form

To help reduce notational clutter let

$$\kappa = \eta/\beta - 1 + \delta$$

and

$$\lambda = \eta - 1 + \delta.$$

We can start to write the Hansen's model in the form employed by Blanchard and Kahn by first defining a vector $\mathbf{s}_t^0$ to contain the model's *dynamic*[4] *predetermined* and *non-predetermined* endogenous variables, e.g.

$$\mathbf{s}_t^0 = \left[\begin{array}{cc} \widehat{k}_t & \widehat{c}_t \end{array} \right]'.$$

Second, let the vector $\mathbf{f}_t^0$ contain the model's *static non-predetermined* variables, e.g.

$$\mathbf{f}_t^0 = \left[\begin{array}{ccc} \widehat{y}_t & \widehat{i}_t & \widehat{h}_t \end{array} \right]'.$$

Given the above definitions for κ and λ, the linearized equilibrium conditions for the *dynamic predetermined* and *non-predetermined* variables, i.e. (2.5"') and (3.7"') respectively

$$\eta \widehat{k}_{t+1} = (1-\delta)\widehat{k}_t + (\eta - 1 + \delta)\widehat{i}_t$$

$$-\frac{\eta}{\beta}\hat{c}_t = -\frac{\eta}{\beta}E_t(\hat{c}_{t+1}) + \left(\frac{\eta}{\beta} - 1 + \delta\right) E_t\hat{y}_{t+1} - \left(\frac{\eta}{\beta} - 1 + \delta\right) \hat{k}_{t+1}$$

can be rewritten as

$$\eta \widehat{k}_{t+1} = (1-\delta)\widehat{k}_t + \lambda \widehat{i}_t$$

$$\kappa \hat{k}_{t+1} + \frac{\eta}{\beta}E_t(\hat{c}_{t+1}) - \kappa E_t\hat{y}_{t+1} = -\frac{\eta}{\beta}\hat{c}_t$$

or (given the definitions of $\mathbf{f}_t^0$ of $\mathbf{s}_t^0$) in matrix form as

$$\mathbf{D}E_t\mathbf{s}_{t+1}^0 + \mathbf{F}E_t\mathbf{f}_{t+1}^0 = \mathbf{G}\mathbf{s}_t^0 + \mathbf{H}\mathbf{f}_t^0 \tag{4.1}$$

[3] You will also find this case referred to as "indeterminate" in the literature.

[4] *Dynamic* in this context means that the variable is determined by a difference equation involving leads and/or lags whereas *static* indicates that a variable is determined by other variables with the same time subscript.

where

$$\mathbf{D} = \begin{bmatrix} \eta & 0 \\ \kappa & \eta/\beta \end{bmatrix},$$

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & 0 \\ -\kappa & 0 & 0 \end{bmatrix},$$

$$\mathbf{G} = \begin{bmatrix} 1-\delta & 0 \\ 0 & \eta/\beta \end{bmatrix},$$

and

$$\mathbf{H} = \begin{bmatrix} 0 & \lambda & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Likewise given the definitions for κ and λ, the linearized equilibrium conditions for the *static* non-*predetermined* variables, i.e. (2.2"'), (2.4"') and (3.5"') respectively

$$\widehat{y}_t = \widehat{a}_t + \theta\widehat{k}_t + (1-\theta)\widehat{h}_t$$

$$(\eta/\beta - 1 + \delta)\widehat{y}_t = [(\eta/\beta - 1 + \delta) - \theta(\eta - 1 + \delta)]\,\widehat{c}_t + \theta(\eta - 1 + \delta)\widehat{i}_t$$

$$\widehat{c}_t + \widehat{h}_t = \widehat{y}_t$$

can be rewritten respectively as

$$\widehat{y}_t + (\theta - 1)\widehat{h}_t = \theta\widehat{k}_t + \widehat{a}_t$$

$$\kappa\widehat{y}_t - \theta\lambda\widehat{i}_t = [\kappa - \theta\lambda]\,\widehat{c}_t$$

$$\widehat{y}_t - \widehat{h}_t = \widehat{c}_t$$

or (given the definitions of $\mathbf{f}_t^0$ of $\mathbf{s}_t^0$) in matrix form as

$$\mathbf{A}\mathbf{f}_t^0 = \mathbf{B}\mathbf{s}_t^0 + \mathbf{C}\widehat{a}_t, \tag{4.2}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \theta - 1 \\ \kappa & -\theta\lambda & 0 \\ 1 & 0 & -1 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} \theta & 0 \\ 0 & \kappa - \theta\lambda \\ 0 & 1 \end{bmatrix},$$

and

$$\mathbf{C} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Finally, the linearized equilibrium condition for the exogenous process describing technological progress, i.e. (2.3"')

$$\widehat{a}_t = \rho\widehat{a}_{t-1} + \varepsilon_t$$

can be solved forward[5] by: first rearranging (2.3"') for $\widehat{a}_{t-1}$ and leading the resulting equation by one period and taking expectations of both sides, e.g.

$$\widehat{a}_t = \frac{1}{\rho}E_t\widehat{a}_{t+1} + E_t\varepsilon_{t+1}.$$

Next we can lead the above equation by another period

$$\widehat{a}_{t+1} = \frac{1}{\rho}E_{t+1}\widehat{a}_{t+2} + E_{t+1}\varepsilon_{t+2};$$

and then substitute this back into the previous equation for $\widehat{a}_{t+1}$, e.g.

$$\begin{aligned} \widehat{a}_t &= \frac{1}{\rho}E_t\left(\frac{1}{\rho}E_{t+1}\widehat{a}_{t+2} + E_{t+1}\varepsilon_{t+2}\right) + E_t\varepsilon_{t+1} \\ &= \left(\frac{1}{\rho}\right)^2 E_tE_{t+1}\widehat{a}_{t+2} + E_tE_{t+1}\varepsilon_{t+2} + E_t\varepsilon_{t+1} \\ &= \left(\frac{1}{\rho}\right)^2 E_t\widehat{a}_{t+2} + E_t(\varepsilon_{t+2} + \varepsilon_{t+1}), \text{ where } E_tE_{t+1} = E_t \\ &= \left(\frac{1}{\rho}\right)^2 E_t\widehat{a}_{t+2}, \text{ since } E_t(\varepsilon_{t+2} + \varepsilon_{t+1}) = 0. \end{aligned}$$

Repeated substitution of this last expression by j periods gives

$$\widehat{a}_t = \left(\frac{1}{\rho}\right)^j E_t\widehat{a}_{t+j}.$$

Finally, rearranging for $E_t\widehat{a}_{t+j}$ yields

$$E_t\widehat{a}_{t+j} = \rho^j\widehat{a}_t. \tag{4.3}$$

[5] See the Appendicies Review 2 for further details on solving linear difference equations.

4.3. Transforming the linearized RBC model to the Blanchard-Kahn form

Now that we have written the linearized system in matrix form, we can next express equations (4.1)-(4.3), i.e.

$$\mathbf{D}E_t\mathbf{s}^0_{t+1} + \mathbf{F}E_t\mathbf{f}^0_{t+1} = \mathbf{G}\mathbf{s}^0_t + \mathbf{H}\mathbf{f}^0_t$$

$$\mathbf{A}\mathbf{f}^0_t = \mathbf{B}\mathbf{s}^0_t + \mathbf{C}\widehat{a}_t,$$

$$E_t\widehat{a}_{t+j} = \rho^j\widehat{a}_t$$

in the same form as the structural model described in the Blanchard-Kahn setup described above.

We start by pre-multiplying both sides of (4.2) by $\mathbf{A}^{-1}$

$$\mathbf{f}^0_t = \mathbf{A}^{-1}\mathbf{B}\mathbf{s}^0_t + \mathbf{A}^{-1}\mathbf{C}\widehat{a}_t. \tag{4.4}$$

Substituting (4.4) into (4.1) yields

$$\begin{aligned}
\mathbf{D}E_t\mathbf{s}^0_{t+1}+\mathbf{F}E_t(\mathbf{A}^{-1}\mathbf{B}\mathbf{s}^0_{t+1}+\mathbf{A}^{-1}\mathbf{C}\widehat{a}_{t+1}) &= \mathbf{G}\mathbf{s}^0_t+\mathbf{H}(\mathbf{A}^{-1}\mathbf{B}\mathbf{s}^0_t+\mathbf{A}^{-1}\mathbf{C}\widehat{a}_t) \\
(\mathbf{D}+\mathbf{F}\mathbf{A}^{-1}\mathbf{B})E_t\mathbf{s}^0_{t+1}+\mathbf{F}\mathbf{A}^{-1}\mathbf{C}E_t\widehat{a}_{t+1} &= (\mathbf{G}+\mathbf{H}\mathbf{A}^{-1}\mathbf{B})\mathbf{s}^0_t+\mathbf{H}\mathbf{A}^{-1}\mathbf{C}\widehat{a}_t \\
(\mathbf{D}+\mathbf{F}\mathbf{A}^{-1}\mathbf{B})E_t\mathbf{s}^0_{t+1}+\mathbf{F}\mathbf{A}^{-1}\mathbf{C}\rho\widehat{a}_t &= (\mathbf{G}+\mathbf{H}\mathbf{A}^{-1}\mathbf{B})\mathbf{s}^0_t+\mathbf{H}\mathbf{A}^{-1}\mathbf{C}\widehat{a}_t \\
(\mathbf{D}+\mathbf{F}\mathbf{A}^{-1}\mathbf{B})E_t\mathbf{s}^0_{t+1} &= (\mathbf{G}+\mathbf{H}\mathbf{A}^{-1}\mathbf{B})\mathbf{s}^0_t+(\mathbf{H}\mathbf{A}^{-1}\mathbf{C}-\mathbf{F}\mathbf{A}^{-1}\mathbf{C}\rho)\widehat{a}_t.
\end{aligned} \tag{4.5}$$

Equation (4.5) can be expressed more succinctly as

$$E_t\mathbf{s}^0_{t+1} = \mathbf{K}\mathbf{s}^0_t + \mathbf{L}\widehat{a}_t \tag{4.6}$$

or

$$E_t\begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{c}_{t+1} \end{bmatrix} = \mathbf{K}\begin{bmatrix} \widehat{k}_t \\ \widehat{c}_t \end{bmatrix} + \mathbf{L}\widehat{a}_t$$

where

$$\mathbf{K} = (\mathbf{D}+\mathbf{F}\mathbf{A}^{-1}\mathbf{B})^{-1}(\mathbf{G}+\mathbf{H}\mathbf{A}^{-1}\mathbf{B})$$

and

$$\mathbf{L} = (\mathbf{D}+\mathbf{F}\mathbf{A}^{-1}\mathbf{B})^{-1}(\mathbf{H}\mathbf{A}^{-1}\mathbf{C}-\mathbf{F}\mathbf{A}^{-1}\mathbf{C}\rho).$$

Note that (4.6) is now in exactly the same form as structural model described in the Blanchard-Kahn setup above.

As shown in Appendix Review 3, the coefficient matrix $\mathbf{K}$ can be transformed into Jordan canonical form

$$\mathbf{K} = \mathbf{M}^{-1}\mathbf{N}\mathbf{M} \tag{4.7}$$

where

$$\mathbf{N} = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$$

and

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}.$$

The $\mathbf{K}$ matrix is a $(n+m) \times (n+m)$ matrix, where n refers to the number of *dynamic predetermined* and m to the number of *dynamic non-predetermined* variables. In this model $n = 1$, i.e. $\widehat{k}_t$ and $m = 1$, i.e. $\widehat{c}_t$. The diagonal matrix $\mathbf{N}$ contains the eigenvalues of $\mathbf{K}$, where N_1 and N_2 are the roots or eigenvalues of $\mathbf{K}$. We will show below both analytically and numerically that N_1 is inside and N_2 outside the unit circle, hence satisfying the condition for a unique solution set out in Blanchard and Kahn (*op cit.*)[6].

The columns of $\mathbf{M}^{-1}$ are the eigenvectors of $\mathbf{K}$; M_{11} is $n \times n$, M_{12} is $n \times m$, M_{21} is $m \times n$, and M_{22} is $m \times m$. In addition, let

$$\mathbf{L} = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix},$$

where $\mathbf{L}$ is a $(n+m) \times k$, L_1 is $n \times k$ and L_2 is $m \times k$ where k refers to the number of exogenous variables. In this model $k = 1$ (i.e. ε_t - the innovation to the process driving technology).

4.4. Solving the transformed model

Pre-multiplying (4.6): $E_t\mathbf{s}^0_{t+1} = \mathbf{K}\mathbf{s}^0_t + \mathbf{L}\widehat{a}_t$ by $\mathbf{M}$ gives

$$\begin{aligned} \mathbf{M}E_t\mathbf{s}^0_{t+1} &= \mathbf{MK}\mathbf{s}^0_t + \mathbf{ML}\widehat{a}_t \\ &= \mathbf{MM}^{-1}\mathbf{NM}\mathbf{s}^0_t + \mathbf{ML}\widehat{a}_t, \ \textit{where } \mathbf{M}^{-1}\mathbf{NM} = \mathbf{K} \\ &= \mathbf{NM}\mathbf{s}^0_t + \mathbf{ML}\widehat{a}_t \end{aligned}$$

[6]In contrast to the model analysed here where $n = m = 1$, Woodford (2003) discusses a number of models containing two *non-predetermined* or *jump* variables and no *predetermined* variables. These models are typicially comprised of a New Keynesian Phillips Curve and an Euler equation. Alternative ways of assessing the conditions for stability in this more general setup are contained in his Appendix C (see p 670).

or

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} E_t \begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{c}_{t+1} \end{bmatrix} = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{c}_t \end{bmatrix} + \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \widehat{a}_t$$

or

$$E_t s_{1t+1}^1 = N_1 s_{1t}^1 + Q_1 \widehat{a}_t \tag{4.8}$$

and

$$E_t s_{2t+1}^1 = N_2 s_{2t}^1 + Q_2 \widehat{a}_t, \tag{4.9}$$

where

$$s_{1t}^1 = M_{11}\widehat{k}_t + M_{12}\widehat{c}_t, \tag{4.10}$$

$$Q_1 = M_{11}L_1 + M_{12}L_2, \tag{4.10'}$$

$$s_{2t}^1 = M_{21}\widehat{k}_t + M_{22}\widehat{c}_t, \tag{4.11}$$

and

$$Q_2 = M_{21}L_1 + M_{22}L_2. \tag{4.11'}$$

Since the eigenvalue N_2 lies outside the unit circle, (4.9) can be solved forward. For example, we first rearrange (4.9) for s_{2t}^1

$$s_{2t}^1 = (1/N_2)E_t s_{2t+1}^1 - (Q_2/N_2)\widehat{a}_t.$$

Leading s_{2t}^1 by one-period gives

$$s_{2t+1}^1 = (1/N_2)E_{t+1} s_{2t+2}^1 - (Q_2/N_2)\widehat{a}_{t+1}$$

Substituting s_{2t+1}^1 into s_{2t}^1 gives

$$\begin{aligned} s_{2t}^1 &= (1/N_2)E_t[(1/N_2)E_{t+1}s_{2t+2}^1 - (Q_2/N_2)\widehat{a}_{t+1}] - (Q_2/N_2)\widehat{a}_t \\ &= E_t[(1/N_2^2)E_{t+1}s_{2t+2}^1 - (Q_2/N_2^2)\widehat{a}_{t+1}] - (Q_2/N_2)\widehat{a}_t \\ &= (1/N_2^2)E_tE_{t+1}s_{2t+2}^1 - (Q_2/N_2^2)E_t\widehat{a}_{t+1} - (Q_2/N_2)\widehat{a}_t \\ &= (1/N_2^2)E_ts_{2t+2}^1 - (Q_2/N_2)\left[(1/N_2)E_t\widehat{a}_{t+1} - \widehat{a}_t\right], \text{ where } E_tE_{t+1} = E_t. \end{aligned}$$

Leading s_{2t+1}^1 by one-period gives

$$s_{2t+2}^1 = (1/N_2)E_{t+2}s_{2t+3}^1 - (Q_2/N_2)\widehat{a}_{t+2}.$$

Again, substituting s^1_{2t+2} into s^1_{2t} yields

$$\begin{aligned} s^1_{2t} &= (1/N_2^2)E_t[(1/N_2)E_{t+2}s^1_{2t+3} - (Q_2/N_2)\widehat{a}_{t+2}] - (Q_2/N_2)\left[(1/N_2)E_t\widehat{a}_{t+1} - \widehat{a}_t\right] \\ &= (1/N_2^3)E_t s^1_{2t+3} - E_t(Q_2/N_2^3)\widehat{a}_{t+2} - (Q_2/N_2)\left[(1/N_2)E_t\widehat{a}_{t+1} - \widehat{a}_t\right] \\ &= (1/N_2^3)E_t s^1_{2t+3} - (Q_2/N_2)\left[(1/N_2^2)E_t\widehat{a}_{t+2} + (1/N_2)E_t\widehat{a}_{t+1} - \widehat{a}_t\right]. \end{aligned}$$

As shown in the Appendices (see Review 2), we can next substitute out s^1_{2t+3} and carry on successive forward substitutions to derive after $(\infty - 1)$ iterations the following general formulation

$$s^1_{2t} = \left(\frac{1}{N_2}\right)^{\infty} E_t s^1_{2t+\infty} - (Q_2/N_2)\sum_{j=0}^{\infty}(1/N_2)^j E_t\widehat{a}_{t+j}.$$

Given that $N_2 > 1$, implying that $\left(\frac{1}{N_2}\right)^{\infty} = 0$, and $E_t\widehat{a}_{t+j} = \rho^j\widehat{a}_t$ we can easily solve for the equilibrium value of s^1_{2t} as

$$\begin{aligned} s^1_{2t} &= -(Q_2/N_2)\sum_{j=0}^{\infty}(1/N_2)^j E_t\widehat{a}_{t+j} \\ &= -(Q_2/N_2)\sum_{j=0}^{\infty}(\rho/N_2)^j\widehat{a}_t \\ &= -\left(\frac{Q_2/N_2}{1-\rho/N_2}\right)\widehat{a}_t \\ &= \left(\frac{Q_2}{\rho - N_2}\right)\widehat{a}_t. \end{aligned}$$

Substituting this result into (4.11): $s^1_{2t} = M_{21}\widehat{k}_t + M_{22}\widehat{c}_t$ gives

$$M_{21}\widehat{k}_t + M_{22}\widehat{c}_t = \left(\frac{Q_2}{\rho - N_2}\right)\widehat{a}_t$$

and rearranging for gives $\widehat{c}_t$

$$\widehat{c}_t = -(M_{21}/M_{22})\widehat{k}_t + (1/M_{22})\left(\frac{Q_2}{\rho - N_2}\right)\widehat{a}_t$$

or, more succinctly

$$\widehat{c}_t = S_1\widehat{k}_t + S_2\widehat{a}_t \tag{4.12}$$

where

$$\begin{aligned} S_1 &= -(M_{21}/M_{22}) \\ S_2 &= (1/M_{22})\left(\frac{Q_2}{\rho - N_2}\right). \end{aligned}$$

We can next substitute (4.12): $\widehat{c}_t = S_1\widehat{k}_t + S_2\widehat{a}_t$ into (4.10): $s^1_{1t} = M_{11}\widehat{k}_t + M_{12}\widehat{c}_t$ to solve for s^1_{1t}

$$\begin{aligned} s^1_{1t} &= M_{11}\widehat{k}_t + M_{12}(S_1\widehat{k}_t + S_2\widehat{a}_t) \\ &= (M_{11} + M_{12}S_1)\widehat{k}_t + (M_{12}S_2)\widehat{a}_t. \end{aligned} \tag{4.13}$$

Substituting (4.13) into (4.8): $E_t s^1_{1t+1} = N_1 s^1_{1t} + Q_1\widehat{a}_t$ gives

$$\begin{aligned} E_t\left[(M_{11} + M_{12}S_1)\widehat{k}_{t+1} + (M_{12}S_2)\widehat{a}_{t+1}\right] &= N_1\left[(M_{11} + M_{12}S_1)\widehat{k}_t + (M_{12}S_2)\widehat{a}_t\right] + Q_1\widehat{a}_t \\ (M_{11} + M_{12}S_1)\widehat{k}_{t+1} + (M_{12}S_2)\rho\widehat{a}_t &= N_1(M_{11} + M_{12}S_1)\widehat{k}_t + (Q_1 + N_1M_{12}S_2)\widehat{a}_t \\ (M_{11} + M_{12}S_1)\widehat{k}_{t+1} &= N_1(M_{11} + M_{12}S_1)\widehat{k}_t + (Q_1 + N_1M_{12}S_2 - M_{12}S_2\rho)\widehat{a}_t \end{aligned}$$

or, more succinctly

$$\widehat{k}_{t+1} = S_3\widehat{k}_t + S_4\widehat{a}_t \tag{4.14}$$

where

$$S_3 = N_1$$

and

$$S_4 = \frac{(Q_1 + N_1M_{12}S_2 - M_{12}S_2\rho)}{(M_{11} + M_{12}S_1)}.$$

We can next return to (4.4) and undertake several manipulations to render a more parsimonious representation. First recall (4.4), i.e.

$$\begin{aligned} \mathbf{f}^0_t &= \mathbf{A}^{-1}\mathbf{B}\mathbf{s}^0_t + \mathbf{A}^{-1}\mathbf{C}\widehat{a}_t \\ &= \mathbf{A}^{-1}\mathbf{B}\begin{bmatrix} \widehat{k}_t \\ \widehat{c}_t \end{bmatrix} + \mathbf{A}^{-1}\mathbf{C}\widehat{a}_t. \end{aligned}$$

Substituting (4.12): $\widehat{c}_t = S_1\widehat{k}_t + S_2\widehat{a}_t$ for $\widehat{c}_t$, into (4.4) gives

$$\begin{aligned} \mathbf{f}^0_t &= \mathbf{A}^{-1}\mathbf{B}\begin{bmatrix} \widehat{k}_t \\ S_1\widehat{k}_t + S_2\widehat{a}_t \end{bmatrix} + \mathbf{A}^{-1}\mathbf{C}\widehat{a}_t \\ &= \mathbf{A}^{-1}\mathbf{B}\begin{bmatrix} 1 \\ S_1 \end{bmatrix}\widehat{k}_t + \left\{\mathbf{A}^{-1}\mathbf{C} + \mathbf{A}^{-1}\mathbf{B}\begin{bmatrix} 0 \\ S_2 \end{bmatrix}\right\}\widehat{a}_t \end{aligned}$$

or, more simply

$$\mathbf{f}^0_t = \mathbf{S_5}\widehat{k}_t + \mathbf{S_6}\widehat{a}_t \tag{4.15}$$

where

$$\mathbf{S_5} = \mathbf{A}^{-1}\mathbf{B}\begin{bmatrix} 1 \\ S_1 \end{bmatrix}$$

and

$$\mathbf{S_6} = \left(\mathbf{A}^{-1}\mathbf{C} + \mathbf{A}^{-1}\mathbf{B} \begin{bmatrix} 0 \\ S_2 \end{bmatrix} \right).$$

We are now in a position to write the model's solution for all four *non-predetermined* variables ($\widehat{y}_t$, $\widehat{i}_t$, $\widehat{h}_t$, $\widehat{c}_t$) and two *predetermined* variables ($\widehat{k}_t$, $\widehat{a}_t$) using equations (2.3"'): $\widehat{a}_t = \rho\widehat{a}_{t-1} + \varepsilon_t$; (4.12): $\widehat{c}_t = S_1\widehat{k}_t + S_2\widehat{a}_t$; (4.14): $\widehat{k}_{t+1} = S_3\widehat{k}_t + S_4\widehat{a}_t$; and (4.15): $\mathbf{f}_t^0 = \mathbf{S_5}\widehat{k}_t + \mathbf{S_6}\widehat{a}_t$, e.g.

$$\begin{aligned} \mathbf{s}_{t+1} &= \mathbf{\Pi s}_t + \mathbf{W}\varepsilon_{t+1} \\ \begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{a}_{t+1} \end{bmatrix} &= \begin{bmatrix} S_3 & S_4 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{a}_t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varepsilon_{t+1} \end{aligned} \tag{4.16}$$

and

$$\begin{aligned} \mathbf{f}_t &= \mathbf{U s}_t \\ \begin{bmatrix} \widehat{y}_t \\ \widehat{i}_t \\ \widehat{h}_t \\ \widehat{c}_t \end{bmatrix} &= \begin{bmatrix} \left[S_{5_{(3x1)}}\right] & \left[S_{6_{(3x1)}}\right] \\ S_1 & S_2 \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{a}_t \end{bmatrix}. \end{aligned} \tag{4.17}$$

5. Calibration and Solution

Given the simplicity of the model set-up, to further illustrate its structure and solution we will next calibrate it using the parameters provided in Ireland 1999 (see Table 1, pg. 16).

5.1. Parameter values (model without vector AR(1) errors)

The values for the subjective rate of time preference, β (= 0.99) and the depreciation rate, δ (= 0.025) are those suggested by Hansen (1985). Ireland (1999) reports that "the estimate of $\gamma = 0.0045$ matches steady-state hours worked in the model with average hours worked in the data; the estimate of $A = 6.0952$ does the same for detrended output. The estimate of θ implies that capital's share in production is just slightly less that 25 percent. The estimate of $\eta = 1.0039$ makes the annualized, steady-state growth rate of real, per-capital output in the model equal to 1.57 percent. Finally the estimate of $\sigma = 0.0050$ is of the same order of magnitude used throughout the literature, while the estimate of $\rho = 0.9983$ implies that technology shocks are extremely persistent".

Parameters (Ireland 1999)

$0 < \beta < 1$	=0.99	subjective rate of time preference
$\gamma > 0$	=0.0045	parameter governing linearity of utility in hours
$\eta > 1$	=1.0039	gross rate of labour augmenting technological progress
$0 < \theta < 1$	=0.2342	elasticity of Y_t with respect to the K_t or capital's share
$A > 0$	=6.0952	constant term in the process describing A_t
$0 < \delta < 1$	=0.025	depreciation rate of K_t
$-1 < \rho < 1$	=0.9983	1st-order autoregressive parameter in the A_t process
$\sigma^2 > 0$	=0.00025	error variance in the A_t process

5.2. Analytic and numeric calculations

5.2.1. Steady-state

Given the algebraic relationships derived in Section 3.4 and the parameter values above we now can calculate the numeric steady-state levels for each variable in the system

$$\begin{aligned} a &= A \\ &= 6.0952 \end{aligned}$$

$$\begin{aligned} k &= \left(\frac{\theta}{\eta/\beta - 1 + \delta}\right) y \\ &= 22629 \end{aligned}$$

$$\begin{aligned} i &= \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right] y \\ &= 654 \end{aligned}$$

$$\begin{aligned} c &= \left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\} y \\ &= 3118 \end{aligned}$$

$$\begin{aligned} h &= \left(\frac{1-\theta}{\gamma}\right)\left(1 - \frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right)^{-1} \\ &= 206 \end{aligned}$$

$$\begin{aligned} y &= a^{1/(1-\theta)}\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right)^{\theta/(1-\theta)}\left(\frac{1-\theta}{\gamma}\right)\left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\}^{-1} \\ &= 3772 \end{aligned}$$

5.2.2. Deviations of the stationary variables from their steady-states

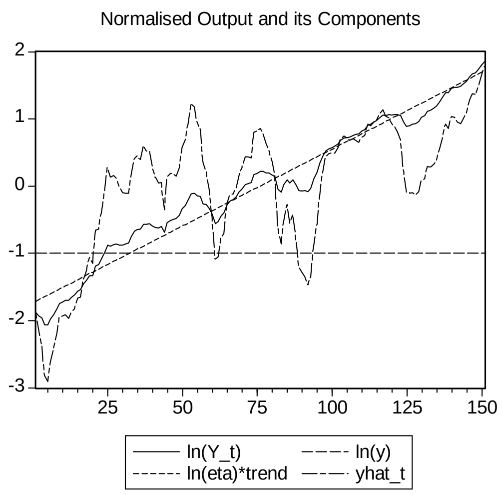
We will see in the next Section that given equations (4.16-4.17) we only require data for Y_t, C_t, I_t, and H_t to estimate the model. Moreover since the resource constraint $Y_t = C_t + I_t$ holds by construction in the data, I_t becomes redundant. Given the results in Sections 3.3 and 3.5 we can transform the levels data to deviations from their steady-states as follows

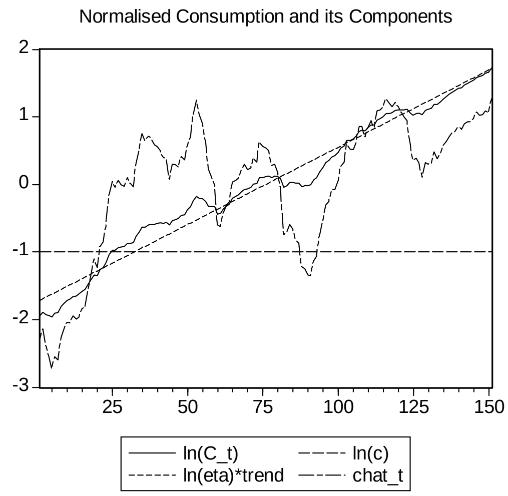
$$\begin{aligned} \widehat{y}_t &= \ln(Y_t) - ln(\eta)trend - ln(y) \\ &= \ln(Y_t) - ln(1.0039)trend - ln(3772) \end{aligned}$$

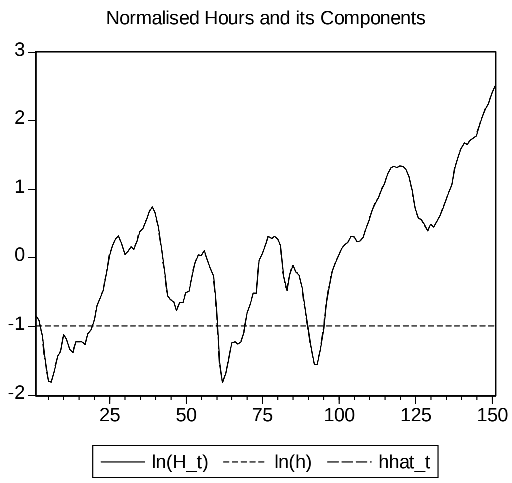
$$\begin{aligned} \widehat{c}_t &= \ln(C_t) - ln(\eta)trend - ln(c) \\ &= \ln(C_t) - ln(1.0039)trend - ln(3118) \end{aligned}$$

$$\begin{aligned} \widehat{h}_t &= \ln(H_t) - ln(h) \\ &= \ln(H_t) - ln(206) \end{aligned}$$

where Y_t is real per capital output defined as $C_t + I_t$, C_t is real (chained 1992$) *per capita* personal consumption expenditure, I_t is real (chained 1992$) *per capita* gross private domestic investment, H_t is hours of wage and salary workers on private, nonfarm payrolls and the population series is the civilian, noninstitutional population, age 16 and over. The normalized plots of the data required to undertake estimation are:







The above graphs were generated using data from 1960.1 to 1997.3 (i.e. the period employed in the Ireland 1999). Prior to plotting, each series was normalized by subtracting its mean and dividing by its standard deviation so that all series for each variable could be viewed simultaneously on the same graph. All data except the population data is seasonally adjusted at an annual rate. The consumption, investment and hours data are from the Reserve Bank of St. Louis, FRED database (http://www.research.stlouisfed.org/fred) and the population data is from the Bureau of Labor Statistics, Establishment Survey (ftp://ftp.bls.gov/pub/special.requests/opt/tableb10.txt).

5.3. Intermediate and final matrix calculations

Using the matrices **A**, **B**, **C**, **D**, **F**, **G** and **H** defined in (4.1-4.17), let's first calculate the various pieces that make up the matrix **K**, i.e. $\mathbf{K} = (\mathbf{D} + \mathbf{F}\mathbf{A}^{-1}\mathbf{B})^{-1}(\mathbf{G} + \mathbf{H}\mathbf{A}^{-1}\mathbf{B})$, see equation 4.6:

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{\theta} & 0 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\sigma^2\lambda} & -\frac{1}{\theta\lambda} & \frac{\kappa(\theta-1)}{\sigma^2\lambda} \\ \frac{1}{\theta} & 0 & -\frac{1}{\theta} \end{bmatrix} = \begin{bmatrix} 4.2699 & 0 & -3.2699 \\ 24.6288 & -147.7458 & -18.8607 \\ 4.2699 & 0 & -4.2699 \end{bmatrix}$$

$$\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} \frac{1}{\theta} & 0 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\sigma^2\lambda} & -\frac{1}{\theta\lambda} & \frac{\kappa(\theta-1)}{\sigma^2\lambda} \\ \frac{1}{\theta} & 0 & -\frac{1}{\theta} \end{bmatrix} \begin{bmatrix} \theta & 0 \\ 0 & \kappa-\theta\lambda \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\theta\lambda} & \frac{\sigma^2\lambda-\kappa}{\sigma^2\lambda} \\ 1 & -\frac{1}{\theta} \end{bmatrix} = \begin{bmatrix} 1 & -3.2699 \\ 5.7681 & -23.6288 \\ 1 & -4.2699 \end{bmatrix}$$

$$\mathbf{F}\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ -\kappa & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\theta\lambda} & \frac{\sigma^2\lambda-\kappa}{\sigma^2\lambda} \\ 1 & -\frac{1}{\theta} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -\kappa & -\frac{\kappa}{\theta}(\theta-1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -0.0390 & 0.1277 \end{bmatrix}$$

$$\begin{aligned}
\mathbf{HA}^{-1}\mathbf{B} &= \begin{bmatrix} 0 & \lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\theta\lambda} & \frac{\sigma^2\lambda-\kappa}{\sigma^2\lambda} \\ 1 & -\frac{1}{\theta} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\kappa}{\theta} & \frac{\sigma^2\lambda-\kappa}{\sigma^2} \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0.1667 & -0.6829 \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{D}+\mathbf{FA}^{-1}\mathbf{B} &= \begin{bmatrix} \eta & 0 \\ \kappa & \eta/\beta \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\kappa & -\frac{\kappa}{\theta}(\theta-1) \end{bmatrix} \\
&= \begin{bmatrix} \eta & 0 \\ 0 & \frac{\eta}{\beta}-\frac{\kappa}{\theta}(\theta-1) \end{bmatrix} \\
&= \begin{bmatrix} -1.0039 & 0 \\ 0 & 1.1417 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
(\mathbf{D}+\mathbf{FA}^{-1}\mathbf{B})^{\mathbf{-1}} &= \begin{bmatrix} \frac{1}{\eta} & 0 \\ 0 & \frac{\beta\theta}{\eta\theta-\beta\kappa(\theta-1)} \end{bmatrix} \\
&= \begin{bmatrix} 0.9961 & 0 \\ 0 & 0.8759 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{G}+\mathbf{HA}^{-1}\mathbf{B} &= \begin{bmatrix} 1-\delta & 0 \\ 0 & \eta/\beta \end{bmatrix} + \begin{bmatrix} \frac{\kappa}{\theta} & \frac{\sigma^2\lambda-\kappa}{\sigma^2} \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1-\delta+\frac{\kappa}{\theta} & \frac{\sigma^2\lambda-\kappa}{\sigma^2} \\ 0 & \frac{\eta}{\beta} \end{bmatrix} \\
&= \begin{bmatrix} 1.1417 & -0.6829 \\ 0 & 1.0140 \end{bmatrix}
\end{aligned}$$

hence

$$\begin{aligned}
\mathbf{K} &= (\mathbf{D}+\mathbf{FA}^{-1}\mathbf{B})^{-\mathbf{1}}(\mathbf{G}+\mathbf{HA}^{-1}\mathbf{B}) \\
&= \begin{bmatrix} \frac{1}{\eta} & 0 \\ 0 & \frac{\beta\theta}{\eta\theta-\beta\kappa(\theta-1)} \end{bmatrix} \begin{bmatrix} 1-\delta+\frac{\kappa}{\theta} & \frac{\sigma^2\lambda-\kappa}{\sigma^2} \\ 0 & \frac{\eta}{\beta} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\kappa+\theta(1-\delta)}{\eta\theta} & \frac{\sigma^2\lambda-\kappa}{\eta\sigma^2} \\ 0 & \frac{\eta\theta}{\eta\theta-\kappa\beta(\theta-1)} \end{bmatrix} \\
&= \begin{bmatrix} 1.1373 & -0.6802 \\ 0 & 0.8882 \end{bmatrix}.
\end{aligned}$$

Substituting for κ and λ (defined at the start of Section 4.2), $\mathbf{K}$ can be rewritten as

$$\mathbf{K} = \begin{bmatrix} K_{11} & K_{12} \\ 0 & K_{22} \end{bmatrix}$$

where

$$\begin{aligned}
K_{11} &= \frac{\eta-\beta(1-\theta)(1-\delta)}{\beta\eta\theta} \\
&= 1.1373
\end{aligned}$$

$$\begin{aligned}
K_{12} &= \frac{\beta\eta\theta^2-\eta+\beta(1-\theta^2)(1-\delta)}{\beta\eta\theta^2} \\
&= -0.6802
\end{aligned}$$

$$\begin{aligned}
K_{22} &= \frac{\eta\theta}{\eta-\beta(1-\theta)(1-\delta)} \\
&= 0.8882.
\end{aligned}$$

Note that expressing $\mathbf{K}$ in this manner will be useful in subsequent calculations reported below. The intermediate calculations that make up the matrix $\mathbf{L}$, i.e. $\mathbf{L} = (\mathbf{D}+\mathbf{FA}^{-1}\mathbf{B})^{-1}(\mathbf{HA}^{-1}\mathbf{C}-\mathbf{FA}^{-1}\mathbf{C}\rho)$, see equation 4.6, include

$$\begin{aligned}
\mathbf{A}^{-1}\mathbf{C} &= \begin{bmatrix} \frac{1}{\theta} & 0 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\sigma^2\lambda} & -\frac{1}{\theta\lambda} & \frac{\kappa(\theta-1)}{\sigma^2\lambda} \\ \frac{1}{\theta} & 0 & -\frac{1}{\theta} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\theta} \\ \frac{\kappa}{\sigma^2\lambda} \\ \frac{1}{\theta} \end{bmatrix} \\
&= \begin{bmatrix} 4.2699 \\ 24.6288 \\ 4.2699 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{HA}^{-1}\mathbf{C} &= \begin{bmatrix} 0 & \lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\theta} \\ \frac{\kappa}{\sigma^2 \lambda} \\ \frac{1}{\theta} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\kappa}{\sigma^2} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 0.7118 \\ 0 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{FA}^{-1}\mathbf{C}\boldsymbol{\rho} &= \begin{bmatrix} 0 & 0 & 0 \\ -\kappa & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\kappa}{\sigma^2} \\ 0 \end{bmatrix} [\boldsymbol{\rho}] \\
&= \begin{bmatrix} 0 \\ -\frac{\kappa\boldsymbol{\rho}}{\theta} \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ -0.1664 \end{bmatrix}
\end{aligned}$$

accordingly

$$\begin{aligned}
\mathbf{L} &= (\mathbf{D} + \mathbf{FA}^{-1}\mathbf{B})^{-\mathbf{1}}(\mathbf{HA}^{-1}\mathbf{C} - \mathbf{FA}^{-1}\mathbf{C}\boldsymbol{\rho}) \\
&= \begin{bmatrix} \frac{1}{\eta} & 0 \\ 0 & \frac{\beta\theta}{\eta\theta - \beta\kappa(\theta-1)} \end{bmatrix} \left(\begin{bmatrix} \frac{\kappa}{\sigma^2} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{\kappa\boldsymbol{\rho}}{\theta} \end{bmatrix} \right) \\
&= \begin{bmatrix} \frac{\kappa}{\eta\theta^2} \\ \frac{\beta\kappa\boldsymbol{\rho}}{\eta\theta - \beta\kappa(\theta-1)} \end{bmatrix} \\
&= \begin{bmatrix} 0.7090 \\ 0.1458 \end{bmatrix}.
\end{aligned}$$

Substituting for κ, as above, $\mathbf{L}$ can be rewritten as

$$\mathbf{L} = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}$$

where

$$\begin{aligned}
L_1 &= \frac{\eta - \beta(1-\delta)}{\beta\eta\theta^2} \\
&= 0.7090
\end{aligned}$$

and

$$L_2 = \frac{\rho\left[\eta - \beta(1-\delta)\right]}{\eta - \beta(1-\theta)(1-\delta)} = 0.1458.$$

The eigenvalues of $\mathbf{K}$ are

$$\begin{aligned} \lambda_r &= K_{11} \\ &= \frac{\eta - \beta(1-\theta)(1-\delta)}{\beta\eta\theta} \\ &= 1.1373 \end{aligned}$$

and

$$\begin{aligned} \lambda_s &= K_{22} \\ &= \frac{\eta\theta}{\eta - \beta(1-\theta)(1-\delta)} \\ &= 0.8882. \end{aligned}$$

As mentioned above, this model has one unstable and one stable root. It can be easily verified, using the theoretical restriction on the parameters, instead of the numeric values assumed above that $K_{11} > 1$ and $0 < K_{22} < 1$. For example using the parameter range defined in Section 2.6

$$\begin{aligned} K_{11} - 1 &= \frac{\eta - \beta(1-\theta)(1-\delta) - \beta\eta\theta}{\beta\eta\theta} \\ &= \frac{\eta(1-\beta\theta) - \beta(1-\theta)(1-\delta)}{\beta\eta\theta} \\ &= \frac{(1-\beta\theta)(\eta-1-\delta) + (1-\beta\theta)(1-\delta) - \beta(1-\theta)(1-\delta)}{\beta\eta\theta} \\ &= \frac{(1-\beta\theta)(\eta-1-\delta) + \left[(1-\beta\theta) - \beta(1-\theta)\right](1-\delta)}{\beta\eta\theta} \\ &= \frac{(1-\beta\theta)(\eta-1-\delta) + (1-\beta)(1-\delta)}{\beta\eta\theta} > 0. \end{aligned}$$

In other words $K_{11} > 1$. Moreover it immediately follows that

$$K_{22} = \frac{\eta\theta}{\eta - \beta(1-\theta)(1-\delta)} > 0$$

and

$$\begin{aligned} K_{22} - 1 &= \frac{\eta\theta}{\eta - \beta(1-\theta)(1-\delta)} \\ &= \frac{\eta\theta - \eta + \beta(1-\theta)(1-\delta)}{\eta - \beta(1-\theta)(1-\delta)} \\ &= \frac{\eta(\theta - 1) + \beta(1-\theta)(1-\delta)}{\eta - \beta(1-\theta)(1-\delta)} \\ &= \frac{\beta(1-\theta)(1-\delta-\eta/\beta)}{\eta - \beta(1-\theta)(1-\delta)} < 0. \end{aligned}$$

In other words, $0 < K_{22} < 1$. Using the tools discussed in Appendix Review 3 we can next find the eigenvectors corresponding to these eigenvalues, e.g.

$$\mathbf{v}^r = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$\begin{aligned} \mathbf{v}^s &= \begin{bmatrix} 1 \\ \frac{(K_{22}-K_{11})}{K_{12}} \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ \frac{\left(\frac{\eta\theta}{\eta-\beta(1-\theta)(1-\delta)} - \frac{\eta-\beta(1-\theta)(1-\delta)}{\beta\eta\theta}\right)}{\frac{\beta\eta\theta^2-\eta+\beta(1-\theta^2)(1-\delta)}{\beta\eta\theta^2}} \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0.3662 \end{bmatrix}. \end{aligned}$$

Recall that in (4.7) we transformed $\mathbf{K}$ into its Jordan canonical form, i.e.

$$\mathbf{K} = \mathbf{M}^{-1}\mathbf{N}\mathbf{M}$$

where

$$\mathbf{N} = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$$

and

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}.$$

Since the diagonal elements of $\mathbf{N}$ are the eigenvalues of $\mathbf{K}$ with N_1 inside and N_2 outside the unit circle and the columns of $\mathbf{M}^{-1}$ are the eigenvectors of $\mathbf{K}$, we have

$$\begin{aligned} N_1 &= K_{22} \\ &= 0.8882 \end{aligned}$$

$$N_2 = K_{11} = 1.1373$$

$$M_{11} = 0$$

$$M_{12} = \frac{K_{12}}{(K_{22} - K_{11})} = 2.7306$$

$$M_{21} = 1$$

and

$$M_{22} = -\frac{K_{12}}{(K_{22} - K_{11})} = -2.7306$$

where, $\mathbf{M} = \mathbf{I}\mathbf{V}^{-1}$ and $\mathbf{V}=\begin{bmatrix} \mathbf{v^r} & \mathbf{v^s} \end{bmatrix}$.

From (4.10') and (4.11') we can find the **Q**'s, i.e.

$$Q_1 = M_{11}L_1 + M_{12}L_2 = \frac{K_{12}L_2}{(K_{22} - K_{11})} = 0.3981$$

and

$$Q_2 = M_{21}L_1 + M_{22}L_2 = L_1 - \frac{K_{12}L_2}{(K_{22} - K_{11})} = 0.3109.$$

Finally, we can find the **S**'s defined in equations (4.12-4.15):

$$S_1 = -(M_{21}/M_{22}) = \frac{(K_{22} - K_{11})}{K_{12}} = 0.3662$$

$$S_2 = (1/M_{22})\left(\frac{Q_2}{\rho - N_2}\right) = \frac{1}{-\frac{K_{12}}{(K_{22}-K_{11})}}\left(\frac{L_1 - \frac{K_{12}L_2}{(K_{22}-K_{11})}}{(\rho - K_{11})}\right) = \frac{(K_{22} - K_{11})L_1 - K_{12}L_2}{K_{12}(K_{11} - \rho)} = 0.8193$$

$$\begin{aligned} S_3 &= N_1 = K_{22} \\ &= 0.882 \end{aligned}$$

$$\begin{aligned} S_4 &= \frac{(Q_1 + N_1 M_{12} S_2 - M_{12} S_2 \rho)}{(M_{11} + M_{12} S_1)} \\ &= \frac{\left(\frac{K_{12} L_2}{(K_{22}-K_{11})}\right) + \left(K_{22} \frac{K_{12}}{(K_{22}-K_{11})} \frac{(K_{22}-K_{11})L_1 - K_{12}L_2}{K_{12}(K_{11}-\rho)}\right) - \frac{K_{12}}{(K_{22}-K_{11})} \frac{(K_{22}-K_{11})L_1 - K_{12}L_2}{K_{12}(K_{11}-\rho)} \rho}{\left(\frac{K_{12}}{(K_{22}-K_{11})} \frac{(K_{22}-K_{11})}{K_{12}}\right)} \\ &= \frac{K_{12} L_2}{(K_{22} - K_{11})} + \frac{(K_{22} - \rho)[(K_{22} - K_{11}) L_1 - K_{12} L_2]}{(K_{22} - K_{11})(K_{11} - \rho)} \\ &= 0.1517 \end{aligned}$$

$$\begin{aligned} \mathbf{S_5} &= \mathbf{A}^{-1} \mathbf{B} \begin{bmatrix} 1 \\ S_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \frac{1}{\theta}(\theta - 1) \\ \frac{\kappa}{\theta \lambda} & \frac{\sigma^2 \lambda - \kappa}{\sigma^2 \lambda} \\ 1 & -\frac{1}{\theta} \end{bmatrix} \begin{bmatrix} 1 \\ S_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 - \left(\frac{1-\theta}{\theta}\right) S_1 \\ S_1 + \left[\frac{\eta/\beta - 1 + \delta}{\sigma^2(\eta - 1 + \delta)}\right] (\theta - S_1) \\ 1 - \left(\frac{1}{\theta}\right) S_1 \end{bmatrix} \\ &= \begin{bmatrix} -0.1973 \\ -2.8840 \\ -0.5635 \end{bmatrix} \end{aligned}$$

and

$$
\begin{aligned}
\mathbf{S_6} &= \mathbf{A}^{-1}\mathbf{C} + \mathbf{A}^{-1}\mathbf{B}\begin{bmatrix} 0 \\ S_2 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\theta} \\ \frac{\kappa}{\sigma^2\lambda} \\ \frac{1}{\theta} \end{bmatrix} + \begin{bmatrix} 1 & \frac{1}{\theta}(\theta-1) \\ \frac{\kappa}{\theta\lambda} & \frac{\sigma^2\lambda-\kappa}{\sigma^2\lambda} \\ 1 & -\frac{1}{\theta} \end{bmatrix}\begin{bmatrix} 0 \\ S_2 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\theta} - \left(\frac{1-\theta}{\theta}\right) S_2 \\ S_2 + \left[\frac{\eta/\beta-1+\delta}{\sigma^2(\eta-1+\delta)}\right](1-S_2) \\ \frac{1}{\theta} - \left(\frac{1}{\theta}\right) S_2 \end{bmatrix} \\
&= \begin{bmatrix} 1.5908 \\ 5.2689 \\ 0.7714 \end{bmatrix}.
\end{aligned}
$$

5.4. Model solution

In the steady-state, i.e. when $\varepsilon_t = 0$ for all $t = 0, 1, 2, ...$the equilibrium conditions in Section 3.3, i.e. (2.2")-(2.5"), (3.5") and (3.7") can be rewritten as follows:

$$0 = y - ak^{\theta}h^{1-\theta}$$

$$0 = a - a$$

$$0 = y - c - i$$

$$0 = (\eta - 1 + \delta)k - i$$

$$0 = \gamma ch - (1-\theta)y$$

and

$$0 = \eta - \beta\left[\theta\left(y/k\right) + 1 - \delta\right].$$

Also recall from Section 3.4 that we solved the equilibrium conditions in Section 3.3 for k, i, c, h assuming that y was in hand and then solved for y, i.e. (3.9)-(3.13)

$$k = \left(\frac{\theta}{\eta/\beta - 1 + \delta}\right) y$$

$$i = \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right] y$$

$$c = \left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\} y$$

$$h = \left(\frac{1-\theta}{\gamma}\right)\left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\}^{-1}$$

$$y = a^{1/(1-\theta)}\left(\frac{\theta}{\eta/\beta - 1 + \delta}\right)^{\theta/(1-\theta)}\left(\frac{1-\theta}{\gamma}\right)\left\{1 - \left[\frac{\theta(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta}\right]\right\}^{-1}.$$

If the algebraic calculations undertaken in Sections 3.3 and 3.4 are correct then substituting the numeric values for k, i, c, h and y that we just calculated will satisfy the equalities (2.2")-(2.5"), (3.5") and (3.7"). This was checked in MATLAB (see Section 7 below) and indeed each of the equilibrium conditions is equal to zero.

Moreover we can also check that the equilibrium conditions hold in the linearized system in Section 3.5, i.e. (2.2"')-(2.5"'), (3.5"') and (3.7"')

$$0 = \widehat{y}_t - \widehat{a}_t - \theta\widehat{k}_t + (1-\theta)\widehat{h}_t$$

$$0 = \widehat{a}_t - \rho\widehat{a}_{t-1}$$

$$0 = (\eta/\beta - 1 + \delta)\left(\widehat{y}_t - \widehat{c}_t\right) + \theta(\eta - 1 + \delta)\left(\widehat{c}_t - \widehat{i}_t\right)$$

$$0 = \eta\widehat{k}_{t+1} - (1-\delta)\widehat{k}_t - (\eta - 1 + \delta)\widehat{i}$$

$$0 = \widehat{c}_t + \widehat{h}_t - \widehat{y}_t$$

$$0 = \frac{\eta}{\beta}(\hat{c}_t - \hat{c}_{t+1}) + \left(\frac{\eta}{\beta} - 1 + \delta\right)\left(\hat{y}_{t+1} - \hat{k}_{t+1}\right).$$

To undertake this check we can make use of the model's solution, i.e. (4.16) and (4.17)

$$\mathbf{s}_{t+1} = \mathbf{\Pi}\mathbf{s}_t + \mathbf{W}\varepsilon_{t+1}$$

$$\mathbf{f}_t = \mathbf{U}\mathbf{s}_t.$$

For example starting from the steady-state in time period t let's temporarily shock ε_t by σ. Given that we are starting in the steady-state, $\mathbf{s}_{t-1} = 0$ therefore $\mathbf{s}_t = \mathbf{W}\varepsilon_t$ and $\mathbf{f}_t = \mathbf{U}\mathbf{s}_t$. So in time period t, $\widehat{k}_t = \mathbf{s}_t(1,1)$, $\widehat{a}_t = \mathbf{s}_t(2,1)$ and $\hat{y}_t = \mathbf{f}_t(1,1)$, $\widehat{i}_t = \mathbf{f}_t(2,1)$, $\widehat{h}_t = \mathbf{f}_t(3,1)$, $\widehat{c}_t = \mathbf{f}_t(4,1)$. In period $t+1$ after the shock, $\mathbf{s}_{t+1} = \mathbf{\Pi}\mathbf{s}_t$ and $\mathbf{f}_{t+1} = \mathbf{U}\mathbf{s}_{t+1}$. Accordingly $\widehat{k}_{t+1} = \mathbf{s}_{t+1}(1,1)$, $\widehat{a}_{t+1} = \mathbf{s}_{t+1}(2,1)$ and $\hat{y}_{t+1} = \mathbf{f}_{t+1}(1,1)$, $\widehat{i}_{t+1} = \mathbf{f}_{t+1}(2,1)$, $\widehat{h}_{t+1} = \mathbf{f}_{t+1}(3,1)$, $\widehat{c}_{t+1} = \mathbf{f}_{t+1}(4,1)$; and so on for $t+2, ...t+T$. To check our algebra leading to equations (4.16) and (4.17) and the numeric calculations we have undertaken above we can substitute the elements of the $\mathbf{s}$ and $\mathbf{f}$ vectors for periods t and $t+1$ into (2.2"', 2.3"', 2.4"', 2.5"', 3.5"' and 3.7"') to confirm that the equilibrium condition is satisfied. This was carried out in MATLAB (see Section 7 below) and indeed each of the linearized equilibrium conditions listed above is equal to zero.

To further illustrate the effects of the temporary shock to ε_t, given our solution in (4.16) and (4.17) and the parameter values on page 22, we can assign numeric values to the system matricies in (4.16) and (4.17), e.g.

$$\begin{aligned}
\mathbf{s}_{t+1} &= \mathbf{\Pi s}_t + \mathbf{W}\varepsilon_{t+1} \\
\begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{a}_{t+1} \end{bmatrix} &= \begin{bmatrix} 0.882 & 0.157 \\ 0 & 0.9983 \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{a}_t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varepsilon_{t+1}
\end{aligned} \tag{4.16$'$}$$

and

$$\begin{aligned}
\mathbf{f}_t &= \mathbf{U s}_t \\
\begin{bmatrix} \widehat{y}_t \\ \widehat{i}_t \\ \widehat{h}_t \\ \widehat{c}_t \end{bmatrix} &= \begin{bmatrix} -0.1973 & 1.5908 \\ -2.8840 & 5.2689 \\ -0.5635 & 0.7714 \\ 0.3662 & 0.8193 \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{a}_t \end{bmatrix}.
\end{aligned} \tag{4.17$'$}$$

Accordingly in time period t, $\widehat{k}_t = \mathbf{s}_t(1,1)$ and $\widehat{a}_t = \mathbf{s}_t(2,1)$, i.e.

$$\begin{aligned}
\widehat{k}_t &= 0.157\widehat{a}_t \\
\widehat{a}_t &= \sigma.
\end{aligned}$$

and $\hat{y}_t = \mathbf{f}_t(1,1)$, $\widehat{i}_t = \mathbf{f}_t(2,1)$, $\widehat{h}_t = \mathbf{f}_t(3,1)$, and $\widehat{c}_t = \mathbf{f}_t(4,1)$, i.e.

$$\begin{aligned}
\widehat{y}_t &= -0.1973\widehat{k}_t + 1.5908\widehat{a}_t \\
\widehat{i}_t &= -2.8840\widehat{k}_t + 5.2689\widehat{a}_t \\
\widehat{h}_t &= -0.5635\widehat{k}_t + 0.7714\widehat{a}_t \\
\widehat{c}_t &= 0.3662\widehat{k}_t + 0.8193\widehat{a}_t.
\end{aligned}$$

In period $t+1$ after the shock,

$$\mathbf{s}_{t+1} = \mathbf{\Pi}\mathbf{s}_t$$

$$\mathbf{f}_{t+1} = \mathbf{U}\mathbf{s}_{t+1}$$

Accordingly, $\widehat{k}_{t+1} = \mathbf{s}_{t+1}(1,1)$, and $\widehat{a}_{t+1} = \mathbf{s}_{t+1}(2,1)$, i.e.

$$\begin{aligned}
\widehat{k}_{t+1} &= 0.882\widehat{k}_t + 0.157\widehat{a}_t \\
\widehat{a}_{t+1} &= 0.9983\widehat{a}_t
\end{aligned}$$

and $\hat{y}_{t+1} = \mathbf{f}_{t+1}(1,1)$, $\widehat{i}_{t+1} = \mathbf{f}_{t+1}(2,1)$, $\widehat{h}_{t+1} = \mathbf{f}_{t+1}(3,1)$, and $\widehat{c}_{t+1} = \mathbf{f}_{t+1}(4,1)$, i.e.

$$\begin{aligned}
\widehat{y}_{t+1} &= -0.1973\widehat{k}_{t+1} + 1.5908\widehat{a}_{t+1} \\
\widehat{i}_{t+1} &= -2.8840\widehat{k}_{t+1} + 5.2689\widehat{a}_{t+1} \\
\widehat{h}_{t+1} &= -0.5635\widehat{k}_{t+1} + 0.7714\widehat{a}_{t+1} \\
\widehat{c}_{t+1} &= 0.3662\widehat{k}_{t+1} + 0.8193\widehat{a}_{t+1}
\end{aligned}$$

and so on for $t+2, ...t+T$.

6. Estimation

6.1. State space form (without *VAR(1)* errors)

As discussed in Section 5.1 we will only require data, $\{\mathbf{d}_t\}_{t=1}^{T}$ for output, consumption and hours to estimate the model since by construction, investment is the residual between output and consumption whilst the capital stock and technological progress will be treated as unobservables. According we can start by defining our measured data vector for all $t = 0, 1, 2, ...$ as

$$\mathbf{d}_t = \begin{bmatrix} \widehat{y}_t & \widehat{c}_t & \widehat{h}_t \end{bmatrix}'.$$

To distinguish the theoretical model from the empirical model we can rewrite equations (4.16) and (4.17) respectively

$$\mathbf{s}_{t+1} = \mathbf{\Pi s}_t + \mathbf{W}\varepsilon_{t+1}$$

$$\mathbf{f}_t = \mathbf{U s}_t$$

as

$$\mathbf{s}_{t+1} = \mathbf{A s}_t + \mathbf{B}\varepsilon_{t+1} \tag{6.1}$$

$$\mathbf{d}_t = \mathbf{C s}_t \tag{6.2}$$

where $\mathbf{A} = \mathbf{\Pi}$, $\mathbf{B} = \mathbf{W}$, $\mathbf{C}$ is comprised of the first, fourth and third rows of $\mathbf{U}$ respectively[7] and $E\varepsilon_{t+1}^2 = V_1 = \sigma^2$. For example

$$\begin{aligned} \mathbf{s}_{t+1} &= \mathbf{A s}_t + \mathbf{B}\varepsilon_{t+1} \\ \begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{a}_{t+1} \end{bmatrix} &= \begin{bmatrix} S_3 & S_4 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{a}_t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varepsilon_{t+1} \end{aligned}$$

and

$$\begin{aligned} \mathbf{d}_t &= \mathbf{C s}_t \\ \begin{bmatrix} \widehat{y}_t \\ \widehat{c}_t \\ \widehat{h}_t \end{bmatrix} &= \begin{bmatrix} S_5(1,1) & S_6(1,1) \\ S_1 & S_2 \\ S_5(3,1) & S_6(3,1) \end{bmatrix} \begin{bmatrix} \widehat{k}_t \\ \widehat{a}_t \end{bmatrix}. \end{aligned}$$

[7]Recall that in (4.17): $\mathbf{f}_t = \begin{bmatrix} \widehat{y}_t & \widehat{i}_t & \widehat{h}_t & \widehat{c}_t \end{bmatrix}'$. Since we are no longer require $\widehat{i}_t$ and $\mathbf{d}_t = \begin{bmatrix} \widehat{y}_t & \widehat{c}_t & \widehat{h}_t \end{bmatrix}'$ the appropriate rows of $\mathbf{U}$ from (4.17) to use for $\mathbf{C}$ in (6.2) are the first, fourth and third.

6.2. State space form (with *VAR(1)* errors)

As discussed in the introduction, to capture the movements and comovements in the data which the theory cannot explain Ireland develops a hybrid representation of the above model by adding serially correlated errors to the observation or control vector given by (6.2), e.g.

$$\mathbf{d}_t = \mathbf{C}\mathbf{s}_t + \boldsymbol{\upsilon}_t \tag{6.3}$$

and

$$\boldsymbol{\upsilon}_{t+1} = \mathbf{D}\boldsymbol{\upsilon}_t + \boldsymbol{\xi}_{t+1} \tag{6.4}$$

where, $\boldsymbol{\upsilon}_t$ is a $3x1$ vector of serially correlated residuals, $E\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}^{'} = \mathbf{V}_2$ and $E(\varepsilon_{t+1}\boldsymbol{\xi}_{t+1}^{\prime}) = \mathbf{0}_{(1x3)}$. In contrast to Sargent (*op cit.*) and the other authors mentioned in the introduction, Ireland does not assume that $\mathbf{D}$ and $\mathbf{V}_2$ are diagonal, hence allowing for cross variable residual correlation.

Since we have introduced another unobservable variable, $\boldsymbol{\upsilon}_t$ we next need to rewrite (6.1), (6.3) and (6.4) so that they conform to the state space form. First let's define the augmented state or unobservable vector as

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{s}_t \\ \boldsymbol{\upsilon}_t \end{bmatrix}.$$

Let's also define an error vector as

$$\boldsymbol{\eta}_{t+1} = \begin{bmatrix} \mathbf{B}\varepsilon_{t+1} \\ \boldsymbol{\xi}_{t+1} \end{bmatrix}.$$

Given the definitions for $\mathbf{x}_t$ and $\boldsymbol{\eta}_{t+1}$, equations (6.1), (6.3) and (6.4) can be rewritten as

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \boldsymbol{\eta}_{t+1} \tag{6.5}$$

$$\mathbf{d}_t = \mathbf{G}\mathbf{x}_t \tag{6.6}$$

where

$$\mathbf{F} = \begin{bmatrix} \mathbf{A} & 0_{(2x3)} \\ 0_{(3x2)} & \mathbf{D} \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{C} & \mathbf{I}_{(3x3)} \end{bmatrix}$$

$$\boldsymbol{\eta}_{t+1}\tilde{\ } N(\mathbf{0}, \mathbf{Q}), \ \ \mathbf{cov}(\boldsymbol{\eta}_{\mathbf{t+i}}, \boldsymbol{\eta}_{\mathbf{t+j}}) = \mathbf{0}, \ \ \mathbf{i} \neq \mathbf{j}$$

and

$$
\begin{aligned}
\mathbf{Q} &= \mathbf{E}\left(\boldsymbol{\eta}_{t+1}\boldsymbol{\eta}_{t+1}'\right) = E\left(\left[\begin{array}{c} \mathbf{B}\varepsilon_{t+1} \\ \boldsymbol{\xi}_{t+1} \end{array}\right]\left[\begin{array}{cc} \varepsilon_{t+1}'\mathbf{B}' & \boldsymbol{\xi}_{t+1}' \end{array}\right]\right) \\
&= \left[\begin{array}{cc} \mathbf{B}E\left(\varepsilon_{t+1}\varepsilon_{t+1}'\right)\mathbf{B}' & \mathbf{B}E\left(\varepsilon_{t+1}\boldsymbol{\xi}_{t+1}'\right) \\ E\left(\boldsymbol{\xi}_{t+1}\varepsilon_{t+1}'\right)\mathbf{B}' & E\left(\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right) \end{array}\right] \\
&= \left[\begin{array}{cc} \mathbf{B}V_1\mathbf{B}' & 0_{(2x3)} \\ 0_{(3x2)} & \mathbf{V}_2 \end{array}\right]. \qquad (6.7)
\end{aligned}
$$

6.3. Kalman filter

The following application of the Kalman filter to the hybrid Hansen model is a special case of the more general treatment provided in the Appendices (see, Review 4). All the proofs of the main results summarized here are contained in Review 4. To reconcile this presentation/notation with that provided in Ireland we will cross reference the corresponding vector and matrix names used in the Appendices.

6.3.1. State space system

The state space system in Review 4, is given by (13.1.1) and (13.1.2) respectively

$$\boldsymbol{\alpha}_t = \mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t\boldsymbol{\eta}_t$$

$$\mathbf{y}_t = \mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_{t.}$$

Ireland's system (6.5) and (6.6) can be written equivalently as

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t$$

$$\mathbf{d}_t = \mathbf{G}\mathbf{x}_t$$

where the correspondence between Ireland's notation and that used in the Appendices is: $\boldsymbol{\alpha}_t = \mathbf{x}_t$; $\mathbf{T}_t = \mathbf{F}$; $\mathbf{c}_t = \mathbf{0}$; $\mathbf{R}_t = \mathbf{I}$; $\boldsymbol{\eta}_t = \boldsymbol{\eta}_t$; $\mathbf{y}_t = \mathbf{d}_t$; $\mathbf{Z}_t = \mathbf{G}$**;** $\boldsymbol{\varepsilon}_t = \mathbf{0}$ and (6.5) was lagged by one period.

6.3.2. Prediction equations

- **Conditional state mean**, $\widehat{\mathbf{x}}_{t|t-1}$

$$\widehat{\mathbf{x}}_{t|t-1} \equiv \mathbf{F}\widehat{\mathbf{x}}_{t-1} \qquad (6.8)$$

where, $\widehat{\mathbf{x}}_{t|t-1} \equiv E_{t-1}(\mathbf{x}_t)$ and the correspondence between Ireland's notation and that in Review 4 is: $\widehat{\mathbf{x}}_{t|t-1} = \mathbf{a}_{t|t-1}$.

- **Conditional covariance matrix,** $\boldsymbol{\Sigma}_{t|t-1}$

$$\boldsymbol{\Sigma}_{t|t-1} \equiv \mathbf{F}\boldsymbol{\Sigma}_{t-1}\mathbf{F}' + \mathbf{Q} \tag{6.9}$$

where $\boldsymbol{\Sigma}_{t|t-1} \equiv E_{t-1}\left[\left(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right)\left(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right)'\right]$ and the correspondence between Ireland's notation and that in Review 4 is: $\boldsymbol{\Sigma}_{t|t-1} = \mathbf{P}_{t|t-1}$ and $\mathbf{Q} = \mathbf{Q}_t$.

6.3.3. Updating equations

The updating equations using Ireland's notation are given by

$\mathbf{d}_{t\|t-1}$	$= \mathbf{G}\widehat{\mathbf{x}}_{t\|t-1}$	*1-step ahead estimate of* $\mathbf{d}_t$	(6.10)
$\boldsymbol{\Omega}_t$	$= \mathbf{G}\boldsymbol{\Sigma}_{t\|t-1}\mathbf{G}'$	*estimated cov matrix of* $\mathbf{u}_t$	(6.11)
$\mathbf{u}_t$	$= \mathbf{d}_t - \mathbf{d}_{t\|t-1}$	*observation vector estimation error*	(6.12)
$\widehat{\mathbf{x}}_t$	$= \widehat{\mathbf{x}}_{t\|t-1} + \mathbf{K}_t\mathbf{u}_t$;	*updating of the state vector*	(6.13)
$\mathbf{K}_t$	$= \boldsymbol{\Sigma}_{t\|t-1}\mathbf{G}'\boldsymbol{\Omega}_t^{-1}$	*Kalman gain*	(6.14)
$\boldsymbol{\Sigma}_t$	$= \boldsymbol{\Sigma}_{t\|t-1} - \mathbf{K}_t\mathbf{G}\boldsymbol{\Sigma}_{t\|t-1}$	*updating of state-cov matrix*	(6.15)

where the correspondence between Ireland's notation and that in Review 4 is: $\boldsymbol{\Omega}_t = \mathbf{F}_t$; $\mathbf{u}_t = \boldsymbol{\nu}_t$; and $\mathbf{K}_t = \mathbf{K}_t$.

6.4. Maximum likelihood estimation

As shown in the Appendices, the innovations $\{\mathbf{u}_t\}_{t=1}^T$ are used to form the likelihood function for $\{\mathbf{d}_t\}_{t=1}^T$ as

$$\ln L = -\frac{3T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln|\boldsymbol{\Omega}_t| - \frac{1}{2}\sum_{t=1}^{T}\mathbf{u}_t' \ln\boldsymbol{\Omega}_t^{-1}\mathbf{u}_t \tag{6.16}$$

where the error covariance matrix $\boldsymbol{\Omega}_t = \mathbf{G}\boldsymbol{\Sigma}_{t|t-1}\mathbf{G}'$. In the Appendix notation the equivalent expression is given in subsection 13.3.4, i.e. $\mathbf{F}_t = \mathbf{Z}_t\mathbf{P}_{t|t-1}\mathbf{Z}_t' + \mathbf{H}_t$. In Ireland's setup $\mathbf{H}_t = 0$.

6.5. Using MATLAB to maximize the likelihood function

To appreciate how the Kalman Filter is used to maximize the above likelihood function, we will next discuss Ireland's MATLAB programs, **est.m** and **llfn.m** which are briefly described in the Appendices and fully listed in a separate document. The purpose of **est.m** is to maximize (minimize the negative) the log likelihood function for the hybrid Hansen model. When maximizing the log likelihood function, some of the parameters are transformed so that they satisfy theoretical restrictions. The log likelihood function given by equation (6.16) (with transformed parameters) is contained in Ireland's MATLAB function **llfn.m**. The line numbers referred to below correspond with those in the full listing of Ireland's source code. You will need to load the code into the MATLAB editor to cross reference the line numbers.

6.5.1. est.m

```
%load data and choose sample period
global yt ct ht
load ych.dat;
```

Comments: The first command declares that the data *yt*, *ct* and *ht* will be defined globally. This allows separate functions which are not part of the base workspace to access these variables. The next command reads *ych.dat* into MATLAB.

```
% full sample
yt = ych(:,1);
ct = ych(:,2);
ht = ych(:,3);
```

Comments: Lines 34-36 define each vector of data with the associated column in the data file *ych.dat*.

```
% set starting values for full sample
bettr = sqrt(0.99/(1-0.99));
gamtr = 0.0045;
thettr = sqrt(0.20/(1-0.20));
etatr = 0.0051;
delttr = sqrt(0.025/(1-0.025));
atr = 6;
rhotr = 0.9975;
sigtr = 0.0055;
```

dyytr = 1.4187;
dyctr = 0.2251;
dyhtr = -0.4441;
dcytr = 0.0935;
dcctr = 1.0236;
dchtr = -0.0908;
dhytr = 0.7775;
dhctr = 0.3706;
dhhtr = 0.2398;
vyytr = 0.0072;
vcytr = 0.0040;
vcctr = 0.0057;
vhytr = 0.0015;
vhctr = 0.0010;
vhhtr = 0.0000;

Comments: In lines 64-88 starting values are chosen for each of the parameters of the model. The first 8 parameters are defined in Section 2.6. The next 15 parameters are the elements of the **D** and $\mathbf{V_2}$ matrices, defined in Section (6.2) e.g. $\mathbf{D} = \begin{bmatrix} d_{yy} & d_{yc} & d_{yh} \\ d_{cy} & d_{cc} & d_{ch} \\ d_{hy} & d_{hc} & d_{hh} \end{bmatrix}$ & $\mathbf{V_2} = \begin{bmatrix} v_{yy} & v_{cy} & v_{hy} \\ v_{cy} & v_{cc} & v_{hc} \\ v_{hy} & v_{hc} & v_{hh} \end{bmatrix}$. Note that $\mathbf{V}_2$ is symmetric hence we only need to estimate 6 elements. Further note that the transformations used for *bettr*, *thettr* and *delttr* ensure the untransformed parameters are bounded between 0 and unity. For example, $\beta_r = \sqrt{\frac{\beta}{1-\beta}}$ implies that $\beta = \frac{\beta_r^2}{1+\beta_r^2}$ i.e. $0 < \beta < 1$. Undertaking the above transformation of β allows the numerical optimisation procedure to search for any value of β_r between $-\infty$ and $+\infty$ but restricts the estimate of β to be within the bounds suggested by theory.

Finally, a reasonable place to start for the starting values of **D** and $\mathbf{V}_2$ is to: (i) run the model where **D** and $\mathbf{V}_2$ are diagonal; (ii) use the estimated residuals from this model and run an Unrestricted *VAR(1)*; (iii) take the Cholesky decomposition of the variance covariance matrix of errors from the *UVAR(1).* The parameters estimates from step (ii) can be used as the starting values for the elements of **D** and the elements of the matrix resulting from the Cholesky decomposition in step (iii) as the starting values for $\mathbf{V}_2$. The Cholesky decompostion will help to perform the same function for the variances (which must be positive) as the parameter restriction transformations performed above.

For example, the Cholesky decompostion of the starting values from the UVAR variance covariance matrix are equal to $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}'$, where $\widetilde{\mathbf{V}}$ is a lower triangular matrix and $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}'$ is a symmetric positive semidefinite matrix. This transformation allows the numerical optimisation procedure to search for any values (i.e. positive or negative) for the $\widetilde{\mathbf{V}}$ matrix since $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}'$ will ensure that the main diagonal of $\mathbf{V}_2$, is positive. We will return to this issue when discussing the **llfn.m** function below.

% maximize likelihood
bigtheto = [gamtr thettr etatr atr rhotr sigtr dyytr dyctr dyhtr dcytr dcctr dchtr dhytr dhctr dhhtr vyytr vcytr vcctr vhytr vhctr vhhtr]';
options(1) = 1;
options(14) = 10000;
thetstar = fminu('llfn',bigtheto,options);

Comments: Lines 204-215 contain the above block of code. The vector *bigtheto* contains the 21 parameters to be estimated (i.e. all 23 parameters listed above less *bettr* and *delttr* which Ireland fixes prior to estimation, see pg. 1121, of his 2004 paper). FMINU finds the minimum of a function of several variables. In MATLAB 6.5 and 7 this has now been replaced with FMINUNC. OPTIONS(1) controls how much display output is given. A value of 1 is for a tabular display of results. OPTIONS(14) controls the maximum number of function evaluations. The default is 100*number of variables. To estimate the 21 parameters in *thetstar* the MATLAB function FMINUNC requires three inputs: *('llfn',bigtheto,options)*. As mentioned above **llfn.m** is Ireland's MATLAB function which we will discuss next.

6.5.2. llfn.m

The purpose of **llfn.m** is to use the Kalman filter to evaluate the negative log likelihood function for the hybrid Hansen model.
function llfn = llfn(bigthet);
% define variables and parameters
global yt ct ht
bigt = length(yt);
bigthet = real(bigthet);
bettr = sqrt(0.99/(1-0.99));
gamtr = bigthet(1);

```
thettr = bigthet(2);
etatr = bigthet(3);
delttr = sqrt(0.025/(1-0.025));
atr = bigthet(4);
rhotr = bigthet(5);
sigtr = bigthet(6);
dyytr = bigthet(7);
dyctr = bigthet(8);
dyhtr = bigthet(9);
dcytr = bigthet(10);
dcctr = bigthet(11);
dchtr = bigthet(12);
dhytr = bigthet(13);
dhctr = bigthet(14);
dhhtr = bigthet(15);
vyytr = bigthet(16);
vcytr = bigthet(17);
vcctr = bigthet(18);
vhytr = bigthet(19);
vhctr = bigthet(20);
vhhtr = bigthet(21);
```

Comments: The above block of code (i.e. lines 26-56) first defines the function *llfn* which requires *bigthet* as an input. The command *length* returns the length of the vector *yt*. The command *real* constrains the elements of the (*21x1*) *bigthet* vector to be real. Therefore if the estimation produces both real and imaginary numbers, the latter are ignored. Note that *bettr* and *delttr* are not included in the (*21x1*) vector to be estimated since they are constrained to take the values 0.99 and 0.025 respectively.

```
% untransform parameters
beta = bettr^2/(1+bettr^2);
gamma = abs(gamtr);
theta = thettr^2/(1+thettr^2);
eta = 1 + abs(etatr);
delta = delttr^2/(1+delttr^2);
```

```
a = abs(atr);
rho = rhotr;
sig = abs(sigtr);
```

Comments: In the above block of code (i.e. lines 60-67), the values of the parameters defined in Section 2.6 are constrained to conform with the ranges dictated by the theory, e.g. $0 < \beta(= 0.99) < 1$; $\gamma > 0$; $0 < \theta < 1$; $\eta > 1$; $0 < \delta(= 0.025) < 1$; $a > 0$; $\sigma > 0$. Only ρ is freely estimated.

```
% compute steady-state values
kappa = eta/beta - 1 + delta;
lambda = eta - 1 + delta;
hss = ((1-theta)/gamma)/(1-theta*lambda/kappa);
yss = (a^(1/(1-theta)))*((theta/kappa)^(theta/(1-theta)))*hss;
kss = (theta/kappa)*yss;
iss = (theta*lambda/kappa)*yss;
css = yss - iss;
```

Comments: The above block of code refers to lines 71-83. κ and λ are defined in the text at the start of Section 4.2. The steady-state expressions for h, y, k, i and c are derived in Section 3.4.

```
% compute K coefficients
bigk11 = (eta-beta*(1-theta)*(1-delta))/(beta*eta*theta);
bigk12 = (beta*eta*theta^2-eta+beta*(1-theta^2)*(1-delta))/(beta*eta*theta^2);
bigk22 = eta*theta/(eta-beta*(1-theta)*(1-delta));

% compute L coefficients
bigl1 = (eta-beta*(1-delta))/(beta*eta*theta^2);
bigl2 = rho*(eta-beta*(1-delta))/(eta-beta*(1-theta)*(1-delta));

% form S matrices
bigs1 = (bigk22-bigk11)/bigk12;
bigs2 = ((bigk22-bigk11)*bigl1-bigk12*bigl2)/(bigk12*(bigk11-rho));
bigs3 = bigk22;
bigs4 = bigk12*bigl2/(bigk22-bigk11) + (bigk22-rho)*((bigk22-bigk11)*bigl1-bigk12*
```

bigl2)/((bigk22-bigk11)*(bigk11-rho));

bigs5 = [1 - ((1-theta)/theta)*bigs1; bigs1 + (kappa/(theta^2*lambda))*(theta-bigs1); 1 -(1/theta)*bigs1];

bigs6 = [1/theta - ((1-theta)/theta)*bigs2; bigs2 + (kappa/(theta^2*lambda))*(1-bigs2);1/theta - (1/theta)*bigs2];

Comments: The above block of code refers to lines 87-118. $K_{\mathbf{11}}$, $K_{\mathbf{12}}$, $K_{\mathbf{22}}$, $L_{\mathbf{1}}$, $L_{\mathbf{2}}$ and **S** are defined in terms of the underlying structural parameters in Section 5.3.

```
% form matrices PI, W, and U
bigpi = [bigs3 bigs4; 0 rho];
bigw = [0 ; 1];
bigu = [bigs5 bigs6 ; bigs1 bigs2];
```

Comments: **Π**, **W**, and **U** are defined beneath equations (4.16 and 4.17) in Section 4.4. This code is appears in lines 122-126.

```
% form matrices AX, BX, CX, DX, V1X, and V2X
bigax = bigpi;
bigbx = bigw;
bigcx = [ bigu(1,:) ; bigu(4,:) ; bigu(3,:) ];
bigdx = [ dyytr dyctr dyhtr ; ...
dcytr dcctr dchtr ; ...
dhytr dhctr dhhtr ];
dxeig = eig(bigdx);
dxviol = 0;
if max(abs(dxeig)) > 1
dxviol = 1;
end
bigv1x = sig^2;
bigv2x1 = [ vyytr 0 0 ; vcytr vcctr 0 ; vhytr vhctr vhhtr ];
bigv2x = bigv2x1*bigv2x1';
```

Comments: The above block of code refers to lines 130-154. The matrices **A**, **B**, **C**, are defined in Section 6.1 beneath equations (6.1-6.2) and V_1 and **D** & $\mathbf{V}_2$ are defined in Section 6.2 beneath equations (6.1) and (6.3-6.4) respectively. The program allows for a possible constraint violation which if true can be used help the numerical optimization procedure to maximize the likelihood function. For example, if the absolute value of the largest eigenvalue of **D** is greater than unity (i.e. *max(dxeig)>1*, where *dxeig=eig(bigdx)*) then a constraint violation is tagged (i.e. *dxviol=1*). We will return to discuss of how *dxviol* is used below.

Moreover note that in lines 148-154 the scalar V_1 is constrained to be positive, e.g. *bigv1x = sig^2* as are the diagonal elements of $\mathbf{V}_2$, e.g. *bigv2x = bigv2x1*bigv2x1'*. As discussed in relation to lines 64-88, in **est.m** *bigv2x* should be a symmetric matrix with non-negative values on the diagonal. Finally note that *bigv2x1*bigv2x1'* corresponds to $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}'$ discussed in lines 64-88.

```
% form matrices FX, GX, and QX
bigfx = [ bigax zeros(2,3) ; zeros(3,2) bigdx ];
biggx = [ bigcx eye(3) ];
bigqx = [ bigbx*bigv1x*bigbx' zeros(2,3) ; zeros(3,2) bigv2x ];
```

Comments: The **F**, **G** and **Q** matrices are defined in Section 6.2 beneath equations (6.5-6.6). This code is in lines 158-162.

```
% put data in deviation form
trend = 1:bigt;
ythat = log(yt) - log(eta)*trend' - log(yss);
cthat = log(ct) - log(eta)*trend' - log(css);
hthat = log(ht) - log(hss);
dthat = [ ythat cthat hthat ];
```

Comments: The above calculations are in lines 166-174. We calculated and plotted these in Section 5.2.2.

```
% evaluate negative log likelihood
xt = zeros(5,1);
bigsig1 = inv(eye(25)-kron(bigfx,bigfx))*bigqx(:);
bigsigt = reshape(bigsig1,5,5);
llfn = (3*bigt/2)*log(2*pi);
for t = 1:bigt
```

```
ut = dthat(t,:)' - biggx*xt;
omegt = biggx*bigsigt*biggx';
omeginvt = inv(omegt);
llfn = llfn + (1/2)*(log(det(omegt))+ut'*omeginvt*ut);
bigkt = bigfx*bigsigt*biggx'*omeginvt;
xt = bigfx*xt + bigkt*ut;
bigsigt = bigqx + bigfx*bigsigt*bigfx'- bigfx*bigsigt*biggx'*omeginvt*biggx*bigsigt*bigfx';
end
```

Comments: The above code is in lines 178-201. Outside the *for* loop: (i) the state vector *xt* is initialized as $(5x1)$ matrix of zeros; (ii) the state covariance matrix *bigsig1* is initialized *bigsig1 = inv(eye(25)-kron(bigfx,bigfx))*bigqx(:)* (see Ireland 2004, pg. 1224)[8]; (iii) *bigsigt* uses the MATLAB command *reshape* to reform *bigsig1* to a $(5x5)$ matrix; (iv) *llfn* is the constant part of likelihood function given by equation (6.16), prior to the first summation. The *for* loop runs from 1 to T and the program first calculates $\mathbf{u}_t$, $\mathbf{\Omega}_t$ and $\mathbf{\Omega}_t^{-1}$ defined in (6.16). The three recursions after the likelihood function include the state updating (recursion 2) and modified versions of the Kalman gain and the covariance matrix updating equations (recursions 1 and 3) derived in Review 4 (see 13.4.5 and 13.4.6; see also 6.14 and 6.15 for Ireland's variable names)

$$\widetilde{\mathbf{K}}_t = \mathbf{F}\mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1} \qquad \text{(recursion 1)}$$

$$\begin{aligned} \widehat{\mathbf{x}}_t &= \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\mathbf{u}_t \qquad \text{(recursion 2)} \\ &\equiv \mathbf{F}\widehat{\mathbf{x}}_{t-1} + \mathbf{K}_t\mathbf{u}_t \end{aligned}$$

$$\begin{aligned} \widetilde{\mathbf{\Sigma}}_{t+1|t} &= \mathbf{F}\left(\mathbf{\Sigma}_{t|t-1} - \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{G}\mathbf{\Sigma}_{t|t-1}\right)\mathbf{F}' + \mathbf{Q} \\ &= \mathbf{F}\mathbf{\Sigma}_{t|t-1}\mathbf{F}' - \mathbf{F}\mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{G}\mathbf{\Sigma}_{t|t-1}\mathbf{F}' + \mathbf{Q}. \qquad \text{(recursion 3)} \end{aligned}$$

The second recursion is simply the state updating equation (6.13). To obtain the remaining two expressions first recall the *prediction equations* for the conditional state mean and covariance matrix given equations (6.8) and (6.9) respectively

$$\widehat{\mathbf{x}}_{t|t-1} \equiv \mathbf{F}\widehat{\mathbf{x}}_{t-1}$$

$$\mathbf{\Sigma}_{t|t-1} \equiv \mathbf{F}\mathbf{\Sigma}_{t-1}\mathbf{F}' + \mathbf{Q}$$

[8] For alternative approaches to initialising the Kalman Filter see pg. 88 in Harvey (1991).

and the corresponding *updating* equations given by (6.13) and (6.15) respectively

$$\widehat{\mathbf{x}}_t = \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \mathbf{u_t}$$

$$\mathbf{\Sigma}_t = \mathbf{\Sigma}_{t|t-1} - \mathbf{K}_t \mathbf{G} \mathbf{\Sigma}_{t|t-1}.$$

Substituting the Kalman gain (6.14): $\mathbf{K}_t = \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}$ into (6.13) and (6.15) respectively gives

$$\widehat{\mathbf{x}}_t = \widehat{\mathbf{x}}_{t|t-1} + \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{u_t} \tag{6.13'}$$

$$\mathbf{\Sigma}_t = \mathbf{\Sigma}_{t|t-1} - \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{G}\mathbf{\Sigma}_{t|t-1}. \tag{6.15'}$$

Leading the *prediction equations* (6.8) and (6.9) by one period gives

$$\widehat{\mathbf{x}}_{t+1|t} = \mathbf{F}\widehat{\mathbf{x}}_t \tag{6.8'}$$

$$\mathbf{\Sigma}_{t+1|t} = \mathbf{F}\mathbf{\Sigma}_t\mathbf{F}' + \mathbf{Q}. \tag{6.9'}$$

Next we can substitute (6.13') and (6.15') into (6.8') and (6.9') for $\widehat{\mathbf{x}}_t$ and $\mathbf{\Sigma}_t$ respectively to obtain the one step ahead *state updating* equation which contains the modified Kalman gain $\widetilde{\mathbf{K}}_t$ (i.e. recursion 1)

$$\begin{aligned}\widehat{\mathbf{x}}_{t+1|t} &= \mathbf{F}\left(\widehat{\mathbf{x}}_{t|t-1} + \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{u_t}\right) \\ &= \mathbf{F}\widehat{\mathbf{x}}_{t|t-1} + \underbrace{\mathbf{F}\mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}}_{\widetilde{\mathbf{K}}_t}\mathbf{u_t}\end{aligned} \tag{6.8''}$$

and the one-step ahead *covariance matrix* updating equation (i.e. recursion 3).

$$\begin{aligned}\widetilde{\mathbf{\Sigma}}_{t+1|t} &= \mathbf{F}\left(\mathbf{\Sigma}_{t|t-1} - \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{G}\mathbf{\Sigma}_{t|t-1}\right)\mathbf{F}' + \mathbf{Q} \\ &= \mathbf{F}\mathbf{\Sigma}_{t|t-1}\mathbf{F}' - \mathbf{F}\mathbf{\Sigma}_{t|t-1}\mathbf{G}'\mathbf{\Omega}_t^{-1}\mathbf{G}\mathbf{\Sigma}_{t|t-1}\mathbf{F}' + \mathbf{Q}.\end{aligned} \tag{6.9''}$$

Expression (6.9') is also known as a *Riccati* equation, see Harvey (1990, pg. 106).

```
% penalize constraint violations
if dxviol
llfn = llfn + 1e8;
end
if abs(imag(llfn)) > 0
llfn = real(llfn) + 1e8;
end
```

```
if abs(rhotr) > 1
llfn = llfn + 1e8;
end
```

Comments: In lines 207-223 above a large number (i.e. 1e8) is added to the likelihood function if: (i) *dxviol=1*; (ii) the absolute value of the imaginary part of the likelihood function is greater than zero (i.e. *abs(imag(llfn)) > 0*); (iii) the absolute value of *rhot* is greater than unity (i.e. *abs(rhotr) > 1*). Violation of any one of these constraints would imply that the search algorithm is looking in the wrong direction for a solution to the optimization problem. Adding a large number (since we are minimizing the negative likelihood function) forces the algorithm to reoptimize by moving it away from its current path to a minimum.

6.6. Statistical inference with maximum likelihood estimators

6.6.1. Asymptotic properties

The large sample asymptotic properties of ML can be summarized as follows[9]:

- It is consistent

$$p_{\lim}(\widehat{\boldsymbol{\theta}}_{ML}) = \boldsymbol{\theta} \tag{6.17}$$

- It is asymptotically normally distributed

$$\widehat{\boldsymbol{\theta}} \xrightarrow{a} N\left[\boldsymbol{\theta}, \{I(\boldsymbol{\theta})\}^{-1}\right] \tag{6.18}$$

- It is asymptotically efficient and achieves the Cramér-Rao lower bound for consistent estimators, i.e.

$$\begin{aligned} Asy.Var[\widehat{\boldsymbol{\theta}}_{ML}] &\equiv E\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)' \\ &\equiv \{I(\boldsymbol{\theta})\}^{-1} \\ &\simeq -\left\{E\left[\frac{\partial^2 \ln L}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]\right\}^{-1} \\ &= -\left\{E\left[\left(\frac{\partial \ln L}{\partial\boldsymbol{\theta}}\right)\left(\frac{\partial \ln L}{\partial\boldsymbol{\theta}'}\right)\right]\right\}^{-1} \end{aligned} \tag{6.19}$$

[9] See Stuart *et al.* (1999) for more detail on the underlying theory and associated proofs.

where $\boldsymbol{\theta}$ is the true parameter vector, $I(\boldsymbol{\theta})$ is the information matrix, $\frac{\partial^2 \ln L}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$ is the second derivative or Hessian estimate of the information matrix. Given $\{I(\boldsymbol{\theta})\}^{-1}$, to calculate standard errors we simply take the square root of the main diagonal which contains the parameter variances, e.g.

$$\{I(\boldsymbol{\theta})\}^{-1} = \begin{bmatrix} E\left(\widehat{\theta}_1 - \theta_{01}\right)^2 & \ldots\ldots & E\left(\widehat{\theta}_1 - \theta_{01}\right)\left(\widehat{\theta}_n - \theta_{0n}\right) \\ E\left(\widehat{\theta}_2 - \theta_{02}\right)\left(\widehat{\theta}_1 - \theta_{01}\right) & \ldots\ldots & E\left(\widehat{\theta}_2 - \theta_{02}\right)\left(\widehat{\theta}_n - \theta_{0n}\right) \\ . & & . \\ . & \ldots\ldots & . \\ . & & . \\ E\left(\widehat{\theta}_n - \theta_{0n}\right)\left(\widehat{\theta}_1 - \theta_{01}\right) & \ldots\ldots & E\left(\widehat{\theta} n - \theta_{0n}\right)^2 \end{bmatrix}.$$

6.6.2. Derivation of the information matrix

Given the general density function of a discrete population, $f(y \mid \theta)$, where y is a vector of n independent observations and θ is a single parameter, the likelihood function can be defined as

$$L \equiv L(y_1, y_2, \ldots y_n \mid \theta) = f(y_1 \mid \theta) f(y_2 \mid \theta) \ldots f(y_n \mid \theta). \tag{6.20}$$

Since L is the joint density function of the observations

$$\int L dy_1 \int L dy_2 \ldots \int L dy_n = 1. \tag{6.21}$$

Next assume that the first two derivatives of L with respect to θ exist for all θ**.** If we differentiate both sides of (6.21) with respect to θ we obtain[10]

$$\int \frac{\partial L}{\partial \theta} dy_1 \int \frac{\partial L}{\partial \theta} dy_2 \ldots \int \frac{\partial L}{\partial \theta} dy_n = 0. \tag{6.22}$$

One of the properties of density functions that are *regular* is that

$$E\left(\frac{\partial \ln L}{\partial \theta}\right) \equiv E\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) = 0. \tag{6.23}$$

[10]Note that the ability to interchange the operations of differentiation and integration (i.e. only differentiating under the integral) is directly related to the *regularity conditions* required to establish the properties of maximum likelihood estimators. More specifically, to obtain (6.22) from (6.21) we must assume (i) if the limits of integration are finite that they are independent of θ; (ii) if the limits of integration are infinite, the integral resulting from the interchange is convergent for all θ and its integrand is a continuous function of x and θ (see, e.g. Stuart *et al.* 1999 and Davidson and MacKinnon 1993 for more details).

Multiplying (6.22) by L/L gives

$$\int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L dy_1 \int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L dy_2 ... \int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L dy_n = 0. \tag{6.24}$$

Given (6.23) and (6.24) it follows that

$$E\left(\frac{\partial \ln L}{\partial \theta}\right) = \int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L dy_1 \int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L dy_2 ... \int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L dy_n = 0. \tag{6.25}$$

If we next totally differentiate (6.25) we obtain

$$\begin{aligned} 0 &= \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\frac{\partial L}{\partial \theta} + L\frac{\partial}{\partial \theta}\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\right\} dy_1 \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\frac{\partial L}{\partial \theta} + L\frac{\partial}{\partial \theta}\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\right\} dy_2 \\ & ... \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\frac{\partial L}{\partial \theta} + L\frac{\partial}{\partial \theta}\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\right\} dy_n. \end{aligned}$$

which can be rewritten as

$$\begin{aligned} 0 &= \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)^2 + \frac{\partial^2 \ln L}{\partial \theta^2}\right\} L dy_1 \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)^2 + \frac{\partial^2 \ln L}{\partial \theta^2}\right\} L dy_2 \\ & ... \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)^2 + \frac{\partial^2 \ln L}{\partial \theta^2}\right\} L dy_n. \end{aligned}$$

In other words,

$$E\left(\frac{\partial \ln L}{\partial \theta}\right)^2 = -E\left(\frac{\partial^2 \ln L}{\partial \theta^2}\right). \tag{6.27}$$

The Cramér-Rao theorem states

$$Var\left(\widehat{\theta}\right) \geq \frac{1}{E\left[\left(\frac{\partial \ln L}{\partial \theta}\right)^2\right]} = \frac{-1}{E\left(\frac{\partial^2 \ln L}{\partial \theta^2}\right)} \tag{6.28}$$

where either expression on the right-hand side indicates the minimum variance bound (MVB). When $\boldsymbol{\theta}$ is a vector of parameters the multidimensional equivalent of $E\left(\frac{\partial^2 \ln L}{\partial \theta^2}\right)$ is the symmetric matrix

$$\begin{aligned} I(\boldsymbol{\theta}) &= -E\left[\frac{\partial^2 \ln L}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right] \\ &= -E\begin{bmatrix} \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_1^2} & \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_1 \boldsymbol{\theta}_2} & \cdots & \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_1 \boldsymbol{\theta}_k} \\ \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_2 \boldsymbol{\theta}_1} & \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_2^2} & \cdots & \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_2 \boldsymbol{\theta} k} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_k \boldsymbol{\theta}_1} & \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_k \boldsymbol{\theta}_2} & \cdots & \frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_k^2} \end{bmatrix} \end{aligned}$$

where as mentioned above $I(\boldsymbol{\theta})$ is called the information matrix. The multidimensional representation of the Cramér-Rao theorem states that

$$Var[\widehat{\boldsymbol{\theta}}] - \{I(\boldsymbol{\theta})\}^{-1} \tag{6.29}$$

is a positive semidefinite matrix. Accordingly, for any $\widehat{\theta}_{i,}$ the MVB is given by the ith element on the principal diagonal of $\{I(\boldsymbol{\theta})\}^{-1}$.

6.6.3. Computing the Hessian

As shown above, to obtain standard errors we need to first calculate the Hessian which contains both cross and second partial derivatives of the log likelihood function, lnL with respect to each of the elements of $\boldsymbol{\theta}$. To find these, we can start by calculating the first partial derivatives of lnL. Since this will be undertaken numerically, it will be useful to discuss the link between analytic and numeric derivatives.

First recall the definition of an analytic first derivative in continuous time of a function with one variable

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

From this we can obtain the following numerical discrete approximation

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

for small values of h. Similarly we could define the numerical derivative using a backward point as

$$f'(x) \approx \frac{f(x) - f(x-h)}{h}.$$

Second recall that the analytic second derivative is given by

$$f''(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h}$$

or using backward differences

$$\begin{aligned} f''(x) &\approx \frac{[f(x+h) - f(x)]/h - [f(x) - f(x-h)]/h}{h} \\ &\approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}. \end{aligned}$$

Returning now to the task of calculating the numeric first and second partial derivatives of lnL with respect to the elements of $\boldsymbol{\theta}$ note that we can rearrange a first-order Taylor approximation of the likelihood function

$$\ln L(\theta_i + h) \approx \ln L(\theta_i) + h \frac{\partial \ln L}{\partial \theta_i}$$

to obtain the one-sided finite discrete difference formula (see Judd, 1998, pg.36)

$$\frac{\partial \ln L}{\partial \theta_i} \approx \frac{\ln L(\theta_i + h) - \ln L(\theta_i)}{h}, \tag{6.30}$$

where h is chosen according to $h = \max(\epsilon\theta_i, \epsilon)$, and $\epsilon = 10^{-6}$. This ensures that h is small relative to θ_i and, at the same time, stays away from zero[11]. As above, we can employ backward differences to obtain the second derivative

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \theta_i^2} &\approx \frac{1}{h}\left(\frac{[\ln L(\theta_i + h) - \ln L(\theta_i)] - [\ln L(\theta_i) - \ln L(\theta_i - h)]}{h}\right) \\ &\approx \frac{\ln L(\theta_i + h) - 2\ln L(\theta_i) + \ln L(\theta_i - h)}{h^2}. \end{aligned} \tag{6.31}$$

Next, consider two elements of $\boldsymbol{\theta}$, say θ_1 and θ_2. The cross derivatives are given by

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_2} &\approx \frac{1}{h}\left(\frac{\ln L(\theta_1 + h, \theta_2 + h) - \ln L(\theta_1, \theta_2 + h)}{h} - \frac{\ln L(\theta_1 + h, \theta_2) - \ln L(\theta_1, \theta_2)}{h}\right) \\ &\approx \frac{\ln L(\theta_1 + h, \theta_2 + h) - \ln L(\theta_1, \theta_2 + h) - \ln L(\theta_1 + h, \theta_2) + \ln L(\theta_1, \theta_2)}{h^2} \end{aligned} \tag{6.32}$$

where, $\frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_2} = \frac{\partial\left(\frac{\partial \ln L}{\partial \theta_1}\right)}{\partial \theta_2}$ and h is assumed to be the same for θ_1 and θ_2 (as in Ireland's code, see below).

6.7. Calculating the standard errors using MATLAB

To appreciate how to undertake the above calculations we will next discuss Ireland's MATLAB program, **estse.m**. Standard errors are calculated for the hybrid model over the full sample by using the **estse.m** program. This scripts calls both the **llfn.m** and **llfnse.m** functions. The latter is the same as the former except it is used here not to minimize the negative log likelihood but evaluate it at its maximum based on the untransformed parameters (i.e. those obtained after the constrained parameters have been optimally estimated). Given that there is overlap in structure and commands used in these programs, extra detailed comments will only be added to explain the major differences.

[11]Note that in the MATLAB program discussed below Ireland assumes that the h is the same for all the parameters.

6.7.1. estse.m

Lines 37-81 which

- define global variables
- load the data
- set starting values
- maximize likelihood

follow the syntax and structure of **est.m** and estimate the 21 parameters of the model where some have been transformed (including variances) to enforce consistency with the theory. The variances have been restricted to be positive in lines 148-154 of **llfn.m** (see detailed comments above). The vector of 21 estimated parameters is called *thetstar*.

```
% find standard errors
thetstar = real(thetstar);
beta = bettr^2/(1+bettr^2);
gamma = abs(thetstar(1));
theta = thetstar(2)^2/(1+thetstar(2)^2);
eta = 1 + abs(thetstar(3));
delta = delttr^2/(1+delttr^2);
a = abs(thetstar(4));
rho = thetstar(5);
sig = abs(thetstar(6));
dyy = thetstar(7);
dyc = thetstar(8);
dyh = thetstar(9);
dcy = thetstar(10);
dcc = thetstar(11);
dch = thetstar(12);
dhy = thetstar(13);
dhc = thetstar(14);
dhh = thetstar(15);
```

```
cholv = [ thetstar(16) 0 0 ; ...
thetstar(17:18)' 0 ; ...
thetstar(19:21)' ];
bigv = cholv*cholv';
vyy = sqrt(bigv(1,1));
vcc = sqrt(bigv(2,2));
vhh = sqrt(bigv(3,3));
vyc = bigv(1,2);
vyh = bigv(1,3);
vch = bigv(2,3);
tstar = [ gamma theta eta a rho sig ...
dyy dyc dyh dcy dcc dch dhy dhc dhh ...
vyy vcc vhh vyc vyh vch ]';
scalvec = ones(21,1);
scalinv = inv(diag(scalvec));
tstars = diag(scalvec)*tstar;
fstar = llfnse(tstars);
```

Comments: In lines 127-129 *tstar* is formed from the untransformed elements of *thetstar*. Note that when constructing *tstar*, the last three elements are first constrained to be positive in lines 111-115 to avoid negative variances. Once *tstar* is formed the value of the likelihood function implied by these 21 parameters is calculated in line 134 using the **llnfse.m** function (i.e. *fstar*).

```
eee = 1e-6;
epsmat = eee*eye(21);
hessvec = zeros(21,1);
for i = 1:21
hessvec(i) = llfnse(tstars+epsmat(:,i));
end
hessmat = zeros(21,21);
for i = 1:21
for j = 1:21
hessmat(i,j) = (llfnse(tstars+epsmat(:,i)+epsmat(:,j)) ...
-hessvec(i)-hessvec(j)+fstar)/eee^2;
```

```
end
end
bighx = scalinv*inv(hessmat)*scalinv';
sevec = sqrt(diag(bighx));
```

Comments: Now that *fstar* has been calculated, the Hessian matrix can be constructed. Note the following: (i) *eee* corresponds with the h used in Section 6.6.3; (ii) lines 160 and 161 which calculate *hessmat(i,j)* correspond directly with the numerical derivative given by equation (6.32), e.g. *llfnse(tstars+epsmat(:,i)+epsmat(:,j))*$\equiv$ $\ln L(\theta_1 + h, \theta_2 + h)$; *hessvec(i)-hessvec(j)* $\equiv \ln L(\theta_1, \theta_2 + h) - \ln L(\theta_1 + h, \theta_2)$; *fstar*$\equiv$ $\ln L(\theta_1, \theta_2)$ and *eee*$\equiv h^2$; (iii) as pointed out in equation (6.19) the standard errors are equal to the square root of the diagonal of the Hessian matrix.

The output of running **est.m** and **estse.m** is contained in the following Table (see Ireland (2004), Table 1).

Full sample estimates and standard errors

Parameter	Estimate	Standard error
γ	0.0045	0.0001
θ	0.2292	0.0065
η	1.0051	0.0005
A	5.1847	0.5048
ρ	0.9987	0.0018
σ	0.0056	0.0004
d_{yy}	1.3655	0.1572
d_{yc}	0.3898	0.1402
d_{yh}	−0.4930	0.1342
d_{cy}	0.1380	0.0712
d_{cc}	0.9690	0.0565
d_{ch}	−0.1046	0.0665
d_{hy}	0.7153	0.2123
d_{hc}	0.4605	0.1593
d_{hh}	0.2219	0.1566
v_y	0.0070	0.0013
v_c	0.0069	0.0007
v_h	0.0018	0.0021
v_{yc}	0.00002989	0.00001064
v_{yh}	0.00000903	0.00000722
v_{ch}	0.00001237	0.00000573

7. Impulse Responses and Variance Decompositions

7.1. Impulse Responses

To calculate the response of the model's six variables to a temporary shock to the process governing technology, $\widehat{a}_t = \rho\widehat{a}_{t-1} + \varepsilon_t$ we can make use of the model's solution given by (4.16) and (4.17) respectively

$$\mathbf{s}_{t+1} = \mathbf{\Pi}\mathbf{s}_t + \mathbf{W}\varepsilon_{t+1}$$

$$\mathbf{f}_t = \mathbf{U}\mathbf{s}_t.$$

As in Section 5.4, starting from the steady-state (i.e. $\varepsilon_t = 0$) assume that in time period t there is a temporary innovation in technological progress (i.e. ε increases by σ in time t and returns to 0 in all subsequent periods after the shock). Given that we are starting in the steady-state, $\mathbf{s}_{t-1} = 0$ therefore in time period t, (4.16) and (4.17) are given by

$$\mathbf{s}_t = \mathbf{W}\sigma$$

$$\mathbf{f}_t = \mathbf{U}\mathbf{s}_t.$$

In period $t+1$ after the shock

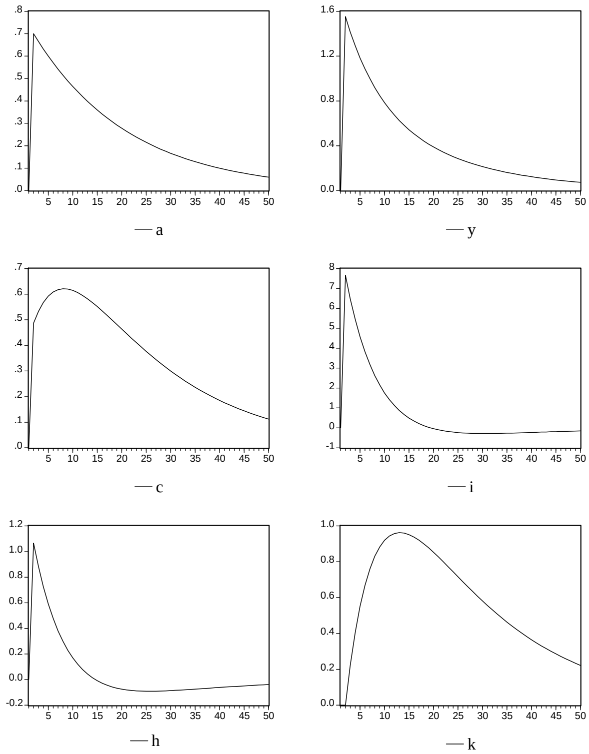
$$\mathbf{s}_{t+1} = \mathbf{\Pi}\mathbf{s}_t$$

$$\mathbf{f}_t = \mathbf{U}\mathbf{s}_t$$

and so for periods $t+2, ...t+T$.

Using the parameter estimates reported above the model's impulse responses are shown below. Note that these can be produced by first running Peter Ireland's MATLAB script called **solv.m** and then **imp.m**. Also the script **check.m** will perform the checks on the levels and log-linearized equilibrium conditions that we undertook in Section 5.4. Given the discussion provided above on the various scripts and functions used to estimate the model parameters and standard errors, these three programs do not require additional comment.

Impulse responses to a temporary technology shock



7.2. Variance decompostions

Recall that the system we are estimating is comprised of equations (6.5) and (6.6) respectively

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \boldsymbol{\eta}_{t+1}$$

$$\mathbf{d}_t = \mathbf{G}\mathbf{x}_t$$

As in Section 6.3.1, we can lag (6.5) by one period to obtain (6.5')

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t.$$

If the process described by (6.5') starts at $t = 1$, we can derive the following representation

$$
\begin{aligned}
\mathbf{x}_1 &= \mathbf{F}\mathbf{x}_0 + \boldsymbol{\eta}_1 \\
\mathbf{x}_2 &= \mathbf{F}\mathbf{x}_1 + \boldsymbol{\eta}_2 = \mathbf{F}\left(\mathbf{F}\mathbf{x}_\mathbf{0} + \boldsymbol{\eta}_1\right) + \boldsymbol{\eta}_2 \\
&= \mathbf{F}^2\mathbf{x}_\mathbf{0} + \mathbf{F}\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 \\
\mathbf{x}_3 &= \mathbf{F}\mathbf{x}_2 + \boldsymbol{\eta}_3 = \mathbf{F}\left(\mathbf{F}^2\mathbf{x}_\mathbf{0} + \mathbf{F}\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2\right) + \boldsymbol{\eta}_3 \\
&= \mathbf{F}^3\mathbf{x}_\mathbf{0} + \mathbf{F}^2\boldsymbol{\eta}_1 + \mathbf{F}\boldsymbol{\eta}_2 + \boldsymbol{\eta}_3 \\
&\quad . \\
&\quad . \\
\mathbf{x}_t &= \mathbf{F}^t\mathbf{x}_\mathbf{0} + \sum_{j=0}^{t-1}\mathbf{F}^j\boldsymbol{\eta}_{t-j}. \qquad (7.1)
\end{aligned}
$$

Accordingly, the vectors $\mathbf{x}_1, ..., \mathbf{x}_t$ and the joint distribution of $\mathbf{x}_1, ..., \mathbf{x}_t$ are uniquely determined by $\mathbf{x}_\mathbf{0}$, $\boldsymbol{\eta}_1 ... \boldsymbol{\eta}_t$. To calculate the forecast error decompositions it is convenient to work with the *MA* representation of (6.5'). In other words it is useful to assume that (6.5') starts not in $t = 1$ but in the infinite past. To derive the *MA* representation we can first write (7.1) as

$$\mathbf{x}_t = \mathbf{F}^{i+1}\mathbf{x}_{t-i-1} + \sum_{j=0}^{i}\mathbf{F}^j\boldsymbol{\eta}_{t-j} \qquad (7.2)$$

where $t-1$ in (7.1) has been replaced by i[12]. If the eigenvalues of $\mathbf{F}$ have modulus less then unity, the sequence $\mathbf{F}^j$, $j = 0, 1, ...$ is absolutely summable (see Lütkepohl, 1991, Appendix A, Section A.9.1), hence the infinite sum $\sum_{i=0}^{\infty}\mathbf{F}^j\boldsymbol{\eta}_{t-j}$ exists in mean square (see Lütkepohl, *op cit.*, Appendix C, Proposition C.7). Moreover since $\mathbf{F}^{i+1}$ converges to zero as $i \to \infty$ we can write the *MA* representation of the *VAR(1)* model given by (6.5') as

$$\mathbf{x}_t = \sum_{j=0}^{\infty}\mathbf{F}^j\boldsymbol{\eta}_{t-j}.$$

or equivalently

$$\mathbf{x}_{t+k} = \sum_{j=0}^{\infty}\mathbf{F}^j\boldsymbol{\eta}_{t+k-j} \qquad (7.3)$$

[12] For example, setting $i = t-1$ in (7.2): $\mathbf{x}_t = \mathbf{F}^{i+1}\mathbf{x}_{t-i-1} + \sum_{j=0}^{i}\mathbf{F}^j\boldsymbol{\eta}_{t-j}$ gives $\mathbf{x}_t = \mathbf{F}^{t-1+1}\mathbf{x}_{t-(t-1)-1} + \sum_{j=0}^{t-1}\mathbf{F}^j\boldsymbol{\eta}_{t-j}$ or (7.1): $\mathbf{x}_t = \mathbf{F}^t\mathbf{x}_\mathbf{0} + \sum_{j=0}^{t-1}\mathbf{F}^j\boldsymbol{\eta}_{t-j}$

where the vector $\mathbf{x}$ is expressed in terms of past and present error innovation vectors $\boldsymbol{\eta}$.

The optimality of the conditional expectation (see Lütkepohl, *op cit.* Section 2.2.2a) implies that

$$\begin{aligned} E_t \mathbf{x}_{t+k} &= \sum_{j=k}^{\infty} \mathbf{F}^j E_t \boldsymbol{\eta}_{t+k-j} \\ &= \sum_{j=k}^{\infty} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j}. \end{aligned} \tag{7.4}$$

Accordingly the forecast error is given by the difference between $\mathbf{x}_{t+k}$ and $E_t\mathbf{x}_{t+k}$

$$\begin{aligned} \mathbf{x}_{t+k} - E_t \mathbf{x}_{t+k} &= \sum_{j=0}^{\infty} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j} - \sum_{j=k}^{\infty} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j} \\ &= \sum_{j=0}^{k-1} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j}. \end{aligned} \tag{7.5}$$

For example, for $k = 1$ subtracting (7.4) from (7.3) gives

$$\mathbf{x}_{t+1} - E_t \mathbf{x}_{t+1} = \mathbf{F}^0 \boldsymbol{\eta}_{t+1} + \mathbf{F}^1 \boldsymbol{\eta}_t + \mathbf{F}^2 \boldsymbol{\eta}_{t-1} + ...\mathbf{F}^{\infty} \boldsymbol{\eta}_{t-\infty-1} - \left(\mathbf{F}^1 \boldsymbol{\eta}_t + \mathbf{F}^2 \boldsymbol{\eta}_{t-1} + ...\mathbf{F}^{\infty} \boldsymbol{\eta}_{t-\infty-1}\right)$$

so for $k = 1$ (7.5) states

$$\mathbf{x}_{t+1} - E_t \mathbf{x}_{t+1} = \mathbf{F}^0 \boldsymbol{\eta}_{t+1}.$$

Given the forecast error, a *k-step* measure of forecast error uncertainty for $\mathbf{x}$ known as the MSE matrix, $\boldsymbol{\Sigma}_k^x$ can be defined as follows

$$\begin{aligned} \boldsymbol{\Sigma}_k^x &= E\left(\mathbf{x}_{t+k} - E_t \mathbf{x}_{t+k}\right)\left(\mathbf{x}_{t+k} - E_t \mathbf{x}_{t+k}\right)' \\ &= E\left[\left(\sum_{j=0}^{k-1} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j}\right)\left(\sum_{j=0}^{k-1} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j}\right)'\right] \\ &= E\left[\left(\mathbf{F}^0 \boldsymbol{\eta}_{t+k} \boldsymbol{\eta}_{t+k}^{'} \mathbf{F}^0\right) + \left(\mathbf{F}^1 \boldsymbol{\eta}_{t+k-1} \boldsymbol{\eta}_{t+k-1}^{'} \mathbf{F}^1\right) + ... \left(\mathbf{F}^{k-1} \boldsymbol{\eta}_{t+1} \boldsymbol{\eta}_{t+1}^{'} \mathbf{F}^{k-1}\right)\right] \\ &= \mathbf{Q}_{t+k} + \mathbf{F}\mathbf{Q}_{t+k-1}\mathbf{F}' + ...\mathbf{F}^{k-1}\mathbf{Q}_{\mathbf{t+1}}\mathbf{F}^{k-1\prime}. \end{aligned} \tag{7.6}$$

Likewise the MSE matrix $\boldsymbol{\Sigma}_k^d$ implied by (6.6): $\mathbf{d}_t = \mathbf{G}\mathbf{x}_t$ can be defined as

$$\begin{aligned} \boldsymbol{\Sigma}_k^d &= E\left(\mathbf{d}_{t+k} - E_t \mathbf{d}_{t+k}\right)\left(\mathbf{d}_{t+k} - E_t \mathbf{d}_{t+k}\right)' \\ &= \mathbf{G}\boldsymbol{\Sigma}_k^x \mathbf{G}'. \end{aligned} \tag{7.7}$$

To derive this expression first note that $\mathbf{d}_t$ can be found by lagging (6.5): $\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \boldsymbol{\eta}_{t+1}$ by one period and substituting it into (6.6): $\mathbf{d}_t = \mathbf{G}\mathbf{x}_t$, e.g.

$$\mathbf{d}_t = \mathbf{G}\mathbf{F}\mathbf{x}_{t-1} + \mathbf{G}\boldsymbol{\eta}_t$$

or equivalently

$$\mathbf{G}\mathbf{x}_t = \mathbf{G}\mathbf{F}\mathbf{x}_{t-1} + \mathbf{G}\boldsymbol{\eta}_t. \tag{7.8}$$

Given the derivations for (7.3) and (7.4) we can write the MA representation of (7.8) as

$$\mathbf{G}\mathbf{x}_{t+k} = \mathbf{G}\sum_{j=0}^{\infty}\mathbf{F}^j\boldsymbol{\eta}_{t+k-j} \tag{7.9}$$

and the conditional expectation of (7.9) as

$$\mathbf{G}E_t\mathbf{x}_{t+k} = \mathbf{G}\sum_{j=k}^{\infty}\mathbf{F}^j\boldsymbol{\eta}_{t+k-j}. \tag{7.10}$$

Accordingly given (7.5) it follows that

$$\begin{aligned}
\boldsymbol{\Sigma}_k^d &= E\left(\mathbf{d}_{t+k} - E_t\mathbf{d}_{t+k}\right)\left(\mathbf{d}_{t+k} - E_t\mathbf{d}_{t+k}\right)' \\
&= E\left(\mathbf{G}\mathbf{x}_{t+k} - E_t\mathbf{G}\mathbf{x}_{t+k}\right)\left(\mathbf{G}\mathbf{x}_{t+k} - E_t\mathbf{G}\mathbf{x}_{t+k}\right)' \\
&= E\left[\left(\mathbf{G}\sum_{j=0}^{k-1}\mathbf{F}^j\boldsymbol{\eta}_{t+k-j}\right)\left(\mathbf{G}\sum_{j=0}^{k-1}\mathbf{F}^j\boldsymbol{\eta}_{t+k-j}\right)'\right] \\
&= \mathbf{G}E\left[\left(\mathbf{F}^0\boldsymbol{\eta}_{t+k}\boldsymbol{\eta}_{t+k}^{'}\mathbf{F}^0\right) + \left(\mathbf{F}^1\boldsymbol{\eta}_{t+k-1}\boldsymbol{\eta}_{t+k-1}^{'}\mathbf{F}^1\right) + ... \left(\mathbf{F}^{k-1}\boldsymbol{\eta}_{t+1}\boldsymbol{\eta}_{t+1}^{'}\mathbf{F}^{k-1}\right)\right]\mathbf{G}' \\
&= \mathbf{G}\left[\mathbf{Q}_{t+k} + \mathbf{F}\mathbf{Q}_{t+k-1}\mathbf{F}' + ...\mathbf{F}^{k-1}\mathbf{Q}_{t+1}\mathbf{F}^{k-1\prime}\right]\mathbf{G}' \\
&= \mathbf{G}\boldsymbol{\Sigma}_k^x\mathbf{G}' \quad \textit{(QED)}.
\end{aligned}$$

7.3. Using Matlab to construct the forecast decompostions and standard errors

The program **vardec.m** uses the parameter estimates and the estimate of Hessian matrix produced by **estse.m** to compute variance decompositions and standard errors for the hybrid Hansen model. Therefore the latter program must be run first and the result matrices retained in the MATLAB workspace. The script **vardec.m** calls the function **vardecfn.m** which contains the formulas used to compute variance decompositions and standard errors.

7.3.1. vardec.m

```
% find covariance matrix of estimated parameters
    bighx = inv(hessmat);

    % compute variance decompositions
    vdec1 = vardecfn(tstar);

    % compute standard errors
    epsmat = 1e-6*eye(21);
    gradmat = zeros(56,21);
    for i = 1:21
    gradmat(:,i) = (vardecfn(tstar+epsmat(:,i))-vdec1)/1e-6;
    end
    vdec2 = gradmat*bighx*gradmat';
    vdec2 = sqrt(diag(vdec2));

    % collect results
    vdec = [ vdec1 vdec2 ];
```

Comments: The output of this script is the ($56x2$) matrix *vdec*. The first column of *vdec* is: *[yvar cvar ivar hvar yvart cvart ivart hvart]'* where *yvar*, *cvar*, *ivar*, and *hvar* are vectors of k-step ahead forecast error variances in output, consumption, investment, and hours worked; *yvart cvart ivart hvart* are percentages due to the technology shock; and k = 1,4,8,12,20,40, and infinity. The variances are all expressed as percentages. The second column of *vdec* gives the standard errors for the corresponding elements of the first column. Note that to calculate the standard errors for the forecast error decompositions, a gradient matrix, i.e. *gradmat(:,i)=(vardecfn(tstar+epsmat(:,i))-vdec1)/1e-6*; is first calculated in line 55. The calculation for *gradmat* is based on the numerical approximation function $f'(x) \approx \frac{f(x+h)-f(x)}{h}$ discussed in Section 6.63. This calculation will be required since the MA parameter matrices of the forecast errors in (7.5) $(\mathbf{x}_{t+k} - E_t\mathbf{x}_{t+k}) = \sum_{j=0}^{k-1} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j}$ are nonlinear functions of the underlying structural parameters ($\boldsymbol{\Theta}$) of the model. For example since $\sum_{j=0}^{k-1} \mathbf{F}^j \boldsymbol{\eta}_{t+k-j} = g(\boldsymbol{\Theta})$, the covariance matrix of the forecast errors is given by $\boldsymbol{\Sigma}_{FE} = \frac{\partial g}{\partial \boldsymbol{\Theta}'} \mathbf{H} \frac{\partial g'}{\partial \boldsymbol{\Theta}}$, where $\mathbf{H}$(=*inv(hessmat)* is the covariance matrix of the structural parameters. In this circumstance the appropriate calculation to obtain $\boldsymbol{\Sigma}_{FE}$ is given in line 59, e.g. *vdec2 = gradmat*bighx*gradmat'* (see, e.g Lütkepohl Section 3.71)[13]. Finally the standard errors for the forecast error decompostion are calculated in line 61, e.g. *vdec2=sqrt(diag(vdec2))*.

7.3.2. vardecfn.m

For given set of parameter estimates this function computes variance decompositions for hybrid Hansen model.

Lines 22-141 which -

- define parameters
- compute steady-state values
- compute K coefficients
- compute L coefficients
- form S matrices
- form matrices PI, W, and U
- form matrices AX, BX, CX, DX, V1X, and V2X
- form matrices FX, GX, and QX

- follow the same syntax and structure as **est.m**.

```
% calculate k-step ahead forecast error variances
vardecfn = zeros(56,1);
bigf2x = eye(5);
bigg2x = zeros(4,5);
bigg2x(1:2,:) = biggx(1:2,:);
bigg2x(3,:) = (yss/iss)*biggx(1,:) - (css/iss)*biggx(2,:);
bigg2x(4,:) = biggx(3,:);
bigsigxk = zeros(5,5);
for k = 1:40
bigsigxk = bigsigxk + bigf2x*bigqx*bigf2x';
```

[13] This is somtimes referred to as the "delta method" (see also Greene (2003) p. 913-914).

```
bigsigyk = bigg2x*bigsigxk*bigg2x';
if k == 1
vardecfn(1) = bigsigyk(1,1);
vardecfn(8) = bigsigyk(2,2);
vardecfn(15) = bigsigyk(3,3);
vardecfn(22) = bigsigyk(4,4);
end
if k == 4
vardecfn(2) = bigsigyk(1,1);
vardecfn(9) = bigsigyk(2,2);
vardecfn(16) = bigsigyk(3,3);
vardecfn(23) = bigsigyk(4,4);
end
if k == 8
vardecfn(3) = bigsigyk(1,1);
vardecfn(10) = bigsigyk(2,2);
vardecfn(17) = bigsigyk(3,3);
vardecfn(24) = bigsigyk(4,4);
end
if k == 12
vardecfn(4) = bigsigyk(1,1);
vardecfn(11) = bigsigyk(2,2);
vardecfn(18) = bigsigyk(3,3);
vardecfn(25) = bigsigyk(4,4);
end
if k == 20
vardecfn(5) = bigsigyk(1,1);
vardecfn(12) = bigsigyk(2,2);
vardecfn(19) = bigsigyk(3,3);
vardecfn(26) = bigsigyk(4,4);
end
if k == 40
vardecfn(6) = bigsigyk(1,1);
vardecfn(13) = bigsigyk(2,2);
vardecfn(20) = bigsigyk(3,3);
```

```
vardecfn(27) = bigsigyk(4,4);
end
bigf2x = bigf2x*bigfx;
end
bigsigx1 = inv(eye(25)-kron(bigfx,bigfx))*bigqx(:);
bigsigx = reshape(bigsigx1,5,5);
bigsigy = bigg2x*bigsigx*bigg2x';
vardecfn(7) = bigsigy(1,1);
vardecfn(14) = bigsigy(2,2);
vardecfn(21) = bigsigy(3,3);
vardecfn(28) = bigsigy(4,4);
```

Comments: The above block code refers to lines 143-217. To calculate the k-step ahead forecast error variances first note lines (158-159) for *bigsigxk* and *bigsigyk* are the analogues of equations (7.6): $\mathbf{\Sigma}_k^x = \mathbf{Q}_{t+k} + \mathbf{F}\mathbf{Q}_{t+k-1}\mathbf{F}' + ... \ \mathbf{F}^{k-1}\mathbf{Q}_{\mathbf{t+1}}\mathbf{F}^{k-1\prime}$ and (7.7): $\mathbf{\Sigma}_k^d = \mathbf{G}\mathbf{\Sigma}_k^x\mathbf{G}'$ (i.e. the MSE matrix for the state and the measurement equations respectively). The program generates k-step ahead forecasts for 7 periods (1, 4, 8, 12, 20, 40 and ∞). Second note that loop around *bigsigxk* and *bigsigyk* runs from 1 to 40. In other words the program generates k-step ahead forecasts for ALL of these periods producing 40 (4x4) *bigsigyk* matrices. Since we only require the periods listed above, the relational operator == is used to select the diagonal elements from the 7 appropriate matrices, i.e. k=1,4,8, etc. The diagonal elements of each of the selected *bigsigyk* matrices contain the k-step ahead forecast error variances for output, consumption, investment and hours respectively. The vector *vardecfn* is (56x1) where its first 7 elements contain the output forecast variances for the 7 forecast horizons. The next 7 elements are taken up by the consumption error variances, etc. The calculation in line (204) i.e. *bigf2x = bigf2x*bigfx* is used to update the **F** matrix in the $\mathbf{\Sigma}_k^x$ expression given by 7.6, e.g. it runs from $\mathbf{F}^0$ to $\mathbf{F}^{k-1}$.

```
% calculate fractions due to technology
bigqxt = [ bigbx*bigv1x*bigbx' zeros(2,3) ; zeros(3,5) ];
bigf2x = eye(5);
bigsigxk = zeros(5,5);
for k = 1:40
bigsigxk = bigsigxk + bigf2x*bigqxt*bigf2x';
```

```
bigsigyk = bigg2x*bigsigxk*bigg2x';
if k == 1
vardecfn(29) = bigsigyk(1,1)/vardecfn(1);
vardecfn(36) = bigsigyk(2,2)/vardecfn(8);
vardecfn(43) = bigsigyk(3,3)/vardecfn(15);
vardecfn(50) = bigsigyk(4,4)/vardecfn(22);
end
```

Comments: To calculate the fraction of the forecast error variance for output, consumption, investment and hours due to technology, *bigsigxk* and *bigsigyk* are recalculated using a **new** *bigqx* matrix, i.e. *bigqxt* in line (221). Recall that $\mathbf{Q}$ defined in (6.7): $\mathbf{Q} = \begin{bmatrix} \mathbf{BV}_1\mathbf{B}' & 0_{(2x3)} \\ 0_{(3x2)} & \mathbf{V}_2 \end{bmatrix}$ is the covariance matrix of the error vector containing the technology shock and the white noise error vector from the VAR. This is the analogue to *bigqx = [bigbx*bigv1x*bigbx' zeros(2,3) ; zeros(3,2) bigv2x]* in the code. In contrast in line (221) *bigqxt = [bigbx*bigv1x*bigbx' zeros(2,3) ; zeros(3,5)]* replaces the (3x3) $\mathbf{V}_2$ and $0_{(3x2)}$ with a $0_{(3x5)}$. In other words, $\mathbf{Q}$ now only contains the error variance of the technology shock, i.e. $E\varepsilon_{t+1}^2 = \mathbf{V}_1 = \sigma^2$ since $\mathbf{B} = [0\ 1]'$. Therefore to calculate the contribution the ratio of the error variance of technology to the total, the new variances calculated in (221) for each of the 7 forecast horizons are divided by the appropriate element of *vardecfn*, e.g. see the block of code under *k==1* above. The structure is the same for the remaining forecast horizons and hence are not listed here. Finally note that the 28 ratios are placed in rows 29-56 of the *vardec* vector.

The output of running **vardec.m** is contained in the following Table (see Ireland (2004), Table 2).

Forecast error variance decompositions

Quarters ahead	Percentage of variance due to technology	Standard error
(A) Output		
1	61.8430	10.5671
4	35.5003	6.9624
8	28.7467	6.7700
12	29.4831	7.5183
20	35.3378	9.1640
40	48.4763	11.0600
∞	89.9399	13.9401
(B) Consumption		
1	31.0978	6.2456
4	32.9700	6.7755
8	35.7260	8.1005
12	39.5799	9.2015
20	48.5522	10.8000
40	65.4138	11.3685
∞	95.7378	6.4410
(C) Investment		
1	44.0529	8.0988
4	25.2808	4.8840
8	18.2636	4.3737
12	17.4674	4.6860
20	18.8007	5.3565
40	21.6648	6.1571
∞	50.6782	32.5590
(D) Hours worked		
1	84.8526	31.3148
4	10.5126	2.1338
8	4.0181	0.9832
12	2.8049	0.8210
20	2.2471	0.7983
40	2.0734	0.8574
∞	2.0609	0.8770

8. Stability Testing

The following testing procedures are from Andrews and Fair (1988).

8.1. Stability of all parameters

We can start by letting $\mathbf{\Theta}^1$ and $\mathbf{\Theta}^2$ denote two vectors of all a model's parameters from two disjoint subsamples. Let's also let $\mathbf{H}^1$ and $\mathbf{H}^2$ represent the covariance matrices associated with these parameter vectors so that asymptotically,

$$\begin{aligned}\mathbf{\Theta}^1 \sim N\left(\mathbf{\Theta}_0^1, \mathbf{H}^1\right) \\ \mathbf{\Theta}^2 \sim N\left(\mathbf{\Theta}_0^2, \mathbf{H}^2\right).\end{aligned} \tag{7.11}$$

To test for the stability of all the estimated parameters, the following likelihood ratio test can be calculated

$$LR = 2\left[\ln L\left(\mathbf{\Theta}^1\right) + \ln L\left(\mathbf{\Theta}^2\right) - \ln L\left(\mathbf{\Theta}\right)\right] \tag{7.12}$$

where $\ln L(\mathbf{\Theta}^1)$, $\ln L(\mathbf{\Theta}^2)$ and $\ln L \ln(\mathbf{\Theta})$ are the values of the maximized log likelihood functions for two sub-samples and the entire sample. Andrews and Fair (*op cit.*) show that this test statistic will be asymptotically distributed as a chi-square random variable with q degrees of freedom under the null hypothesis of parameter stability, where q is equal to the number of estimated parameters.

8.2. Stability of subsets of parameters

To test for the stability of a subset of the parameters contained in $\mathbf{\Theta}^1$ and $\mathbf{\Theta}^2$, i.e. $\mathbf{\Theta}_q^1$ and $\mathbf{\Theta}_q^2$ with associated covariance matrices $\mathbf{H}_q^1$, and $\mathbf{H}_q^2$, Andrews and Fair develop the following Wald test statistic

$$W = \left(\mathbf{\Theta}_q^1 - \mathbf{\Theta}_q^2\right)'\left(\mathbf{H}_q^1 + \mathbf{H}_q^2\right)^{-1}\left(\mathbf{\Theta}_q^1 - \mathbf{\Theta}_q^2\right). \tag{7.13}$$

Andrews and Fair (*op cit.*) show that this test statistic will be asymptotically distributed as a chi-square random variable with q degrees of freedom under the null hypothesis of parameter stability, where q is equal to the number of estimated parameters being tested for stability.

8.3. Using MATLAB to conduct stability tests

The above tests can be calculated by using Peter Ireland's program **stabtest.m** and **stabtst2.m**. The former uses 1973:1 as the break point and the latter uses 1980.1. The output of **stabtest.m**[14] is:

- parameter and standard error estimates from 1948:1-1972:4
- parameter and standard error estimates from 1973:1-2002:2
- LR statistic for stability of **all** parameters, i.e. using equation (7.2)[15]

[14] The output of *stabtst2.m* is the same for the post 1980.1 break point.

[15] Note that β and δ are not included in the tests since they have been imposed and not estimated.

- Wald statistic for stability of **all** parameters, i.e. using equation (7.3)
- Wald statistic for stability of **structural** parameters (i.e.γ, θ, η, a) plus ρ and σ
- Wald statistic for stability of structural parameters, **excluding** average productivity growth, a
- Wald statistic for stability of **other** parameters (i.e. $d_{yy}, d_{yc}, d_{yh}, d_{cy}, d_{cc}, d_{ch}, d_{hy}$, $d_{hc}, d_{hh}, v_{yy}, v_{cy}, v_{cc}, v_{hy}, v_{hc}, v_{hh}$)

Lines 39-305 in **stabtest.m** which

- set global variables
- load data
- define pre-1973 subsample
- set starting values for pre-1973 subsample
- maximize likelihood for pre-1973 subsample
- find standard errors for pre-1973 subsample
- reload data
- define post-1973 subsample
- set starting values for post-1973 subsample
- maximize likelihood for post-1973 subsample
- find standard errors for post-1973 subsample

follow the same syntax and structure as **est.m**, **estse.m** and call the functions **llfn.m**, **llfn2.m**, **llnfse.m**, and **llfnse2.m** to calculate the parameter estimates and standard errors over the two subsamples.

% form test statistics
lrstat = -2*(fstar1+fstar2+2323.5501);
wstat = (tstar1-tstar2)'*inv(bighx1+bighx2)*(tstar1-tstar2);

Comments: Lines 309-311 calculate the LR and Wald tests for all parameters. *fstar1*, *fstar2* and *232.5501* correspond to the value of the likelihood function in subperiod 1, subperiod 2 and the entire sample respectively. *tstar1* and *tstar2* refer to the vector of untransformed parameter estimates in subperiod 1 and 2 respectively. The *bighx1* and *bighx2* matrices are the $(21x21)$ inverse Hessian matrices over the respective subsamples.

```
bighxq1 = bighx1(1:6,1:6);
bighxq2 = bighx2(1:6,1:6);
wstatq = (tstar1(1:6)-tstar2(1:6))' ...
*inv(bighxq1+bighxq2)*(tstar1(1:6)-tstar2(1:6));
```

Comments: Lines 313-317 follow the same structure as lines 309-311 except that only 6 structural parameters are included in the tests (see above). These are the parameters contained in the first 6 rows of the vectors *tstar1* and *tstar2* whereas *bighxq1* and *bighxq2* are $(6x6)$ matrices taken from the first 6 rows and 6 columns of *bighx1* and *bighx2* respectively.

```
bighxx1 = [ bighx1(1:2,1:2) bighx1(1:2,4:6) ; ...
bighx1(4:6,1:2) bighx1(4:6,4:6) ];
bighxx2 = [ bighx2(1:2,1:2) bighx2(1:2,4:6) ; ...
bighx2(4:6,1:2) bighx2(4:6,4:6) ];
tstarx1 = [ tstar1(1:2) ; tstar1(4:6) ];
tstarx2 = [ tstar2(1:2) ; tstar2(4:6) ];
wstatx = (tstarx1-tstarx2)' ...
*inv(bighxx1+bighxx2)*(tstarx1-tstarx2);
```

Comments: Lines 319-330 again follow the same structure as 313-317 except that the relevant elements of the *tstar* vectors and *bighx* matrices have been excluded to accommodate dropping average productivity growth, a from the list.

```
bighxn1 = bighx1(7:21,7:21);
bighxn2 = bighx2(7:21,7:21);
wstatn = (tstar1(7:21)-tstar2(7:21))' ...
*inv(bighxn1+bighxn2)*(tstar1(7:21)-tstar2(7:21));
```

Comments: Lines 332-336 repeat the Wald test for parameter 7-21 (see the list above).

The outputs of running **stabtst1.m** and **stabtst2.m** are contained in the following two Tables (see Ireland (2004), Tables 3 and 4)

Subsample estimates and standard errors

Parameter	Pre-1980 estimate	Standard error	Post-1980 estimate	Standard error
γ	0.0046	0.0001	0.0042	0.0002
θ	0.2190	0.0045	0.2457	0.0077
η	1.0053	0.0005	1.0046	0.0012
A	5.6309	0.2842	5.0535	0.9839
ρ	0.9935	0.0099	0.9966	0.0135
σ	0.0050	0.0015	0.0042	0.0009
d_{yy}	1.2553	0.2013	1.1390	0.1360
d_{yc}	0.1657	0.1558	0.6564	0.2970
d_{yh}	−0.3837	0.1633	−0.5332	0.1092
d_{cy}	0.1216	0.0889	0.0768	0.0999
d_{cc}	0.9065	0.0677	1.1172	0.1409
d_{ch}	−0.1143	0.0601	−0.1479	0.2176
d_{hy}	0.6380	0.2449	0.3722	0.1460
d_{hc}	0.1689	0.2225	0.5337	0.6014
d_{hh}	0.4232	0.2362	0.3562	0.1839
v_y	0.0090	0.0024	0.0056	0.0013
v_c	0.0084	0.0009	0.0051	0.0009
v_h	0.0049	0.0032	0.0013	0.0010
v_{yc}	0.00004982	0.00002080	0.00001542	0.00000774
v_{yh}	0.00002959	0.00003403	0.00000682	0.00000379
v_{ch}	0.00002201	0.00001458	0.00000518	0.00000595

Tests for parameter stability

Stability of all 21 estimated parameters:	$W = 74.5383^{***}$
Stability of the 6 structural parameters:	$W = 15.3546^{**}$
Stability of the 15 remaining parameters:	$W = 54.1925^{***}$

Note: ** and *** indicate significance at the 5% and 1% levels.

.

9. Forecast Accuracy

The following testing procedures are from Diebold and Mariano (1995) and Harvey *et al.* (1997).

9.1. Generating model forecasts

Forecasts can be generated using the system given by (6.5): $\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \boldsymbol{\eta}_{t+1}$ and (6.6): $\mathbf{d}_t = \mathbf{G}\mathbf{x}_t$. For example, the expected value in time t of (6.6) k periods ahead is

$$E_t\mathbf{d}_{t+k} = \mathbf{G}E_t\mathbf{x}_{t+k}. \tag{9.1}$$

For $k = 1$, the conditional expected value of (6.5) is defined as

$$E_t\mathbf{x}_{t+1} \equiv \widehat{\mathbf{x}}_{t+1|t}. \tag{9.2}$$

From (6.8) we know that $\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}\widehat{\mathbf{x}}_{t-1}$. Hence leading (6.8) by one period and substituting for $\widehat{\mathbf{x}}_{t+1|t}$ in (9.2) gives

$$E_t\mathbf{x}_{t+1} = \mathbf{F}\widehat{\mathbf{x}}_t. \tag{9.3}$$

For $k > 1$

$$\begin{aligned} E_t\mathbf{x}_{t+k} &= \mathbf{F}^{k-1}\mathbf{E_t}\mathbf{x}_{t+1} \\ &= \mathbf{F}^{k-1}\mathbf{F}\widehat{\mathbf{x}}_t \\ &= \mathbf{F}^{k}\widehat{\mathbf{x}}_t. \end{aligned} \tag{9.4}$$

Therefore for any $k \geq 1$ it follows from (9.4) that we can rewrite (9.1) as

$$E_t\mathbf{d}_{t+k} = \mathbf{G}\mathbf{F}^{k}\widehat{\mathbf{x}}_t. \tag{9.5}$$

9.2. Evaluating forecast accuracy

First let $\{\widetilde{\varepsilon}_t\}_{t=1}^{T}$ represent a sequence of k-step-ahead forecast errors from the RBC model and let $\{\widetilde{\eta}_t\}_{t=1}^{T}$ represent a sequence of k-step-ahead forecast errors from a competing model (e.g. an alternative theoretical model or an unconstrained *VAR* model). Next let loss functions $g(\widetilde{\varepsilon})$ and $g(\widetilde{\eta})$ denote the mean squared forecast errors

$$g(\widetilde{\varepsilon}) = \frac{1}{T}\sum_{t=1}^{T}\widetilde{\varepsilon}_t^2 \tag{9.6}$$

$$g(\widetilde{\eta}) = \frac{1}{T}\sum_{t=1}^{T}\widetilde{\eta}_t^2 \tag{9.7}$$

and let l_t denote the "loss differential" defined by Diebold and Mariano (*op cit.*) as,

$$l_t = \widetilde{\eta}_t^2 - \widetilde{\varepsilon}_t^2. \tag{9.8}$$

To test the null hypothesis of equal forecast accuracy first calculate the sample mean loss differential, i.e.

$$l = \frac{1}{T}\sum_{t=1}^{T} l_t \tag{9.9}$$

where,

$$\sqrt{T} l \tilde{} N(0, f),$$

$$f = \gamma(0) + 2\sum_{\tau=1}^{k-1}\gamma(\tau),$$

$$\gamma(\tau) = \frac{1}{T}\sum_{t=\tau+1}^{T} l_t l_{t-\tau}.$$

Finally, the proposed test-statistic is given by

$$S = \frac{l}{\sqrt{f/T}} \tag{9.10}$$

which is distributed standard normal assuming that l_t satisfies a number of regularity assumptions, e.g. covariance stationarity, short memory and the existence of moments that ensure the applicability of the central limit theorem.

9.3. Bias corrected S statistic

Diebold and Mariano (*op cit.*) show via Monte-Carlo simulations that the normal distribution can be a poor approximation of the S test's finite null distribution. For example, they show that the S test can have the wrong size, which leads to the null being rejected too often. Harvey *et al.* (1997) have developed a bias corrected S test statistic based on the t-distribution in an effort to improve its small sample properties, e.g.

$$S_c = \sqrt{\frac{T + 1 - 2k + k(k-1)/T}{T}} S \tag{9.11}$$

where S_c depends on l_t being independent and normally distributed.

9.4. Using MATLAB to conduct forecast accuracy tests

The MATLAB code for calculating the S test is provided in the script **fork.m** which calls the function **forkfn.m**. Moreover **fork.m** requires *tstarmat* and *tstrmatd*, which are the matrices of estimated coefficients for the two versions of the RBC model generated by **estseq.m** and **estseqd.m**. The structure of these programs is as follows:

9.4.1. fork.m

fork.m: Generates k-step-ahead forecasts from four competing models which start using the models estimated over the sample period 1948:1-1984:4 and ends using the model estimated over the sample period 1948:1-2002:2. These models include:

- the hybrid Hansen model, where the matrices $\mathbf{D}$ and $\mathbf{V}_2$ are unconstrained
- the hybrid Hansen model, where the matrices $\mathbf{D}$ and $\mathbf{V}_2$ are constrained to be diagonal
- an unconstrained *VAR(1)* model
- an unconstrained *VAR(2)* model

The output of **fork.m** includes:

- forecasts of $[\ln(y_t)\ \log(c_t)\ \ln(i_t)\ \ln(h_t)]$ from the unconstrained model
- forecasts of $[\ln(y_t)\ \log(c_t)\ \ln(i_t)\ \ln(h_t)]$ from the constrained model
- forecasts of $[\ln(y_t)\ \log(c_t)\ \ln(i_t)\ \ln(h_t)]$ from the *VAR(1)*
- forecasts of $[\ln(y_t)\ \log(c_t)\ \ln(i_t)\ \ln(h_t)]$ from the *VAR(2)*
- root mean squared errors, unconstrained model
- root mean squared errors, constrained model
- root mean squared errors, *VAR(1)*
- root mean squared errors, *VAR(2)*
- mean loss differential and standard error, unconstrained model, using the *VAR(1)* as a benchmark

- mean loss differential and standard error, constrained model, using the *VAR(1)* as a benchmark
- mean loss differential and standard error, unconstrained model, using the *VAR(2)* as a benchmark
- mean loss differential and standard error, constrained model, using the *VAR(2)* as a benchmark

Given the discussion provided above on the various scripts and functions used to estimate the model parameters and standard errors; calculate forecast error decompositions, etc. this program should be quite straightforward.

9.4.2. forkfn.m

This function uses the Kalman filter to generate k-step-ahead forecasts for output, consumption, investment, and hours worked from the hybrid Hansen model.

Lines 25-162 which

- define variables and parameters
- compute steady-state values
- compute K coefficients
- compute L coefficients
- form S matrices
- form matrices AX, BX, CX, DX, V1X, and V2X
- form matrices FX, GX, and QX
- put data in deviation form

follows the same syntax and structure as **est.m**.

% generate forecasts of transformed variables
xt = zeros(5,1);
bigsig1 = inv(eye(25)-kron(bigfx,bigfx))*bigqx(:);

```
bigsigt = reshape(bigsig1,5,5);
for t = 1:bigt
ut = dthat(t,:)' - biggx*xt;
omegt = biggx*bigsigt*biggx';
omeginvt = inv(omegt);
bigkt = bigfx*bigsigt*biggx'*omeginvt;
xt = bigfx*xt + bigkt*ut;
bigsigt = bigqx + bigfx*bigsigt*bigfx' ...
- bigfx*bigsigt*biggx'*omeginvt*biggx*bigsigt*bigfx';
end
ytk = biggx*(bigfx^(k-1))*xt;
```

Comments:

Probably the only part of this code requiring extra comment is lines (164-189). Note that forecasts are generated using the Kalman filter via the 3 recursions discussed above, i.e. the modified Kalman gain: *bigkt =bigfx*bigsigt*biggx'* omeginvt*; the state updating: *xt = bigfx*xt + bigkt*ut*; and the covariance matrix updating: *bigsigt = bigqx +bigfx*bigsigt *bigfx'- bigfx*bigsigt*biggx'*omeginvt*biggx*bigsigt*bigfx'* equations. Finally *ytk = biggx*(bigfx^(k-1))*xt is the* analogue to equation (8.5): $E_t\mathbf{d}_{t+k} = \mathbf{G}\mathbf{F}^k\widehat{\mathbf{x}}_t$ since $\mathbf{F}^k\widehat{\mathbf{x}}_t \equiv \mathbf{F}^{k-1}E_t\widehat{\mathbf{x}}_{t+1.}$

The output of running **fork.m** is contained in the following Table (see Ireland (2004), Table 5).

Forecast Accuracy, 1985:1-2002:2

Quarters Ahead	1	2	3	4
(A) Output				
RMSE: Hybrid	0.8319	1.6810	2.5706	3.4501
RMSE: Diagonal	0.6935	1.2201	1.7117	2.1613
RMSE: VAR(1)	1.0163	1.8602	2.6015	3.2472
RMSE: VAR(2)	0.9490	1.9783	2.9944	3.8845
S: Hybrid vs. VAR(1)	4.2047***	1.7496*	0.2110	−0.9795
S: Diagonal vs. VAR(1)	3.9323***	2.3712**	1.7474*	1.4071
S: Hybrid vs. VAR(2)	2.1937**	1.5608	1.1815	0.8642
S: Diagonal vs. VAR(2)	2.8708***	2.1497**	1.7631*	1.5245
(B) Consumption				
RMSE: Hybrid	0.5371	0.8849	1.2361	1.6208
RMSE: Diagonal	0.4781	0.7022	0.9093	1.1688
RMSE: VAR(1)	0.5554	0.8963	1.2110	1.5308
RMSE: VAR(2)	0.5854	1.0228	1.4828	1.9111
S: Hybrid vs. VAR(1)	0.9207	0.2572	−0.2818	−0.6440
S: Diagonal vs. VAR(1)	2.3321**	1.9189*	1.4721	1.1254
S: Hybrid vs. VAR(2)	1.1583	1.4070	1.2813	1.0203
S: Diagonal vs. VAR(2)	2.7518***	2.3575**	1.9652**	1.6083
(C) Investment				
RMSE: Hybrid	3.2441	5.8452	8.8561	11.9088
RMSE: Diagonal	3.1803	4.6807	6.1267	7.4359
RMSE: VAR(1)	4.0511	6.7590	9.1619	11.2436
RMSE: VAR(2)	3.4950	6.6361	9.9190	12.8333
S: Hybrid vs. VAR(1)	4.2287***	2.0408**	0.5264	−0.9300
S: Diagonal vs. VAR(1)	3.0864***	2.3256**	1.7907*	1.4927
S: Hybrid vs. VAR(2)	1.2140	1.3146	0.9559	0.6083
S: Diagonal vs. VAR(2)	1.0464	1.7969*	1.6306	1.4754
(D) Hours worked				
RMSE: Hybrid	0.5345	1.2973	2.2239	3.2436
RMSE: Diagonal	0.5663	1.0753	1.5589	2.0097
RMSE: VAR(1)	0.7439	1.4743	2.2072	2.9236
RMSE: VAR(2)	0.4458	1.1096	1.9253	2.7838
S: Hybrid vs. VAR(1)	4.3000***	1.5516	−0.0953	−1.1794
S: Diagonal vs. VAR(1)	3.5831***	2.5072**	2.2060**	2.0702**
S: Hybrid vs. VAR(2)	−4.4089***	−2.6781***	−2.1450**	−1.9447*
S: Diagonal vs. VAR(2)	−3.4761***	0.3892	1.7386*	1.8918*

Note: *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Appendices

10. Review 1: Alternative Methods of Log-Linearizing

10.1. Method 1: Taylor approximation

The first order Taylor polynomial of any nonlinear function $f(x_t)$ is given by

$$f(x_t) = f(x) + f'(x_t)_{|x_t=x}(x_t - x) + R(x_t). \tag{10.1}$$

where x is the value of x_t at which the approximation is taken. If we allow

$$f(x_t) = e^{x_t} \tag{10.2}$$

and evaluate this at $x_t = 0$ then (10.1) can be written as

$$\begin{aligned} f(x_t) &= e^0 + e^0(x_t - 0) + R(x_t) \\ &= 1 + x_t + R(x_t) \\ &\simeq 1 + x_t \text{ since } \lim_{x_t \to x} R(x_t) \simeq 0. \end{aligned} \tag{10.3}$$

Let x be the steady-state value of a variable x_t. Therefore the log deviation of x_t from its steady-state value, i.e. $\hat{x}_t$, can be defined as $\hat{x}_t = \ln(x_t/x)$. Accordingly,

$$\begin{aligned} e^{\hat{x}_t} &= e^{\ln(x_t/x)} \\ &= x_t/x. \end{aligned} \tag{10.4}$$

The function $e^{\hat{x}_t}$ can also be approximated with a first order Taylor Polynomial around the steady-state value of $\hat{x}_t$, i.e. $\hat{x}_t = 0$, since there are no fluctuations. Let's now redefine $f(.)$ in (10.2) as $f(\hat{x}_t) = e^{\widehat{x}_t}$. As above the first order expansion at $\hat{x}_t = 0$ can be written as

$$\begin{aligned} f(\hat{x}_t) &= e^0 + e^0(\widehat{x}_t - 0) + R(\widehat{x}_t) \\ &= 1 + \widehat{x}_t + R(\widehat{x}_t) \\ &\simeq 1 + \widehat{x}_t \text{ since } \lim_{\widehat{x}_t \to \widehat{x}} R(\widehat{x}_t) \simeq 0. \end{aligned}$$

or

$$e^{\widehat{x}_t} \simeq 1 + \widehat{x}_t \tag{10.5}$$

since $f(\hat{x}_t) = e^{\widehat{x}_t}$. Recall from (10.4) that $e^{\hat{x}_t} \simeq x_t/x$. Substituting for $e^{\hat{x}_t}$ in (10.5) and solving for x_t gives

$$x_t \simeq x(1 + \hat{x}_t).$$

If we next rearrange (10.4) for x_t, i.e. $x_t = xe^{\hat{x}_t}$, we can re-express the above relation equivalently as

$$xe^{\hat{x}_t} \simeq x(1 + \hat{x}_t). \tag{10.6}$$

Consider the following simple Cobb-Douglas production function,

$$y_t = x_t^{\alpha}. \tag{10.7}$$

Given the above discussion, this function can be written equivalently as

$$ye^{\widehat{y}_t} = \left(xe^{\widehat{x}_t}\right)^{\alpha}. \tag{10.8}$$

The first order Taylor series expansion of the both sides of this equation gives

$$\begin{aligned} y(1+\widehat{y}_t) &= x^{\alpha}(1+\alpha\widehat{x}_t) \\ y + y\widehat{y}_t &= x^{\alpha} + x^{\alpha}\alpha\widehat{x}_t \\ \widehat{y}_t &= \alpha\widehat{x}_t. \end{aligned} \tag{10.9}$$

where $\widehat{y}_t = \ln(y_t/y)$ and $\widehat{x}_t = \ln(x_t/x)$.

10.2. Method 2: Total differential of $lnf(x)$

The second method relies on first taking the natural logarithm of $f(x)$ and then taking the total differential of the resulting function and evaluating it at the steady-state. So if we again consider the Cobb-Douglas function in (10.7) we can log-linearize this as follows

$$\begin{aligned} lny_t &= \alpha \ln x \\ d[\ln y_t] &= d[\alpha \ln x] \\ \frac{1}{y}dy_t &= \alpha\frac{1}{x}dx_t \\ \widehat{y}_t &= \alpha\widehat{x}_t. \end{aligned} \tag{10.10}$$

where $\widehat{y}_t = \frac{1}{y}dy_t \cong \ln(y_t/y)$ and $\widehat{x}_t = \frac{1}{y}dx_t \cong \ln(x_t/x)$.

11. Review 2: Alternative Methods of Solving First-Order Difference Equations

11.1. Backward iteration

Consider the following first order autonomous nonhomogeneous linear difference equation relating current Y to its own lag and the parameters a and λ

$$Y_t = a + \lambda Y_{t-1}. \tag{11.1}$$

The equation is autonomous since it is independent of t[16] and nonhomogeneous since $Y_t - \lambda Y_{t-1} \neq 0$. The one period lag of (11.1) is given by

$$Y_{t-1} = a + \lambda Y_{t-2}. \tag{11.2}$$

Substituting (11.2) into (11.1) gives

$$\begin{aligned} Y_t &= a + \lambda(a + \lambda Y_{t-2}) \\ &= a + a\lambda + \lambda^2 Y_{t-2}. \end{aligned} \tag{11.3}$$

The two period lag of (11.1) is

$$Y_{t-2} = a + \lambda Y_{t-3}. \tag{11.4}$$

Substituting (11.4) into (11.3) gives

$$\begin{aligned} Y_t &= a + a\lambda + \lambda^2(a + \lambda Y_{t-3}) \\ &= a + a\lambda + a\lambda^2 + \lambda^3 Y_{t-3}. \end{aligned} \tag{11.5}$$

Next we can substitute out Y_{t-3} and carry on successive backward substitutions to derive after $T-1$ iterations the following general formulation

$$Y_t = a \sum_{i=0}^{T-1} \lambda^i + \lambda^T Y_{t-T}. \tag{11.6}$$

Concentrating on the summation $S = a \sum_{i=0}^{T-1} \lambda^i = a(1 + \lambda + \lambda^2 + ...\lambda^{T-1})$ for the moment, let's first lead S by multiplying it by λ and then subtract the product $S\lambda$ from

[16] In other words, t is not an independent argument on the right hand side of (11.1).

S, e.g.

$$\begin{aligned} S(1-\lambda) &= a\sum_{i=0}^{T-1}\lambda^i - \left(a\sum_{i=0}^{T-1}\lambda^i\lambda\right) \\ &= a\sum_{i=0}^{T-1}\lambda^i - a\sum_{i=0}^{T}\lambda^{i+1} \\ &= a\sum_{i=0}^{T-1}\lambda^i - a\sum_{i=1}^{T}\lambda^i \qquad (11.7) \\ &= a(1+\lambda+\lambda^2+...\lambda^{T-1}) - a(\lambda+\lambda^2+...\lambda^T) \\ &= a - a\lambda^T. \end{aligned}$$

Dividing both sides of (11.7) by $(1-\lambda)$ gives

$$S = \frac{a}{(1-\lambda)} - \frac{a\lambda^T}{(1-\lambda)}. \qquad (11.8)$$

In equilibrium, $Y_t = Y_{t-1}$, accordingly we can suppress time subscripts and write the difference equation given by (11.1) as

$$\begin{aligned} y &= a + \lambda y \qquad (11.9) \\ &= \frac{a}{1-\lambda} \end{aligned}$$

where y (represents the equilibrium value of Y). Therefore in equilibrium S in (11.8) becomes

$$S = y - y\lambda^T. \qquad (11.10)$$

We can now substitute (11.10) into (11.6): $Y_t = a\sum_{i=0}^{T-1}\lambda^i + \lambda^T Y_{t-T}$ to find the general solution to the difference equation given by (11.1)

$$\begin{aligned} Y_t &= y - y\lambda^T + \lambda^T Y_{t-T} \qquad (11.11) \\ &= y + \lambda^T(Y_{t-T} - y). \end{aligned}$$

To find the numeric solution to (11.1) requires that we know an initial or starting value Y_0. If Y_0 is exactly t periods before Y_t then $Y_t = y + \lambda^t(Y_0 - y)$. Thus we can see that the general solution is comprised of the sum of two components, i.e. $y = \frac{a}{1-\lambda}$ and $\lambda^t(Y_0 - y)$ which are know as the *particular integral* and *complementary function* respectively. As mentioned above, the former represents the intertemporal equilibrium level of Y_t whereas the latter is the deviations of the time path from that equilibrium. If $0 < \lambda < 1$ as $t \to \infty$, $Y_t \to y$ and the solution is stable or convergent.

11.2. Forward iteration

Consider again the first order difference equation set out in (11.1)

$$Y_t = a + \lambda Y_{t-1}. \tag{11.12}$$

Solving for Y_{t-1} gives

$$Y_{t-1} = -\frac{a}{\lambda} + \frac{1}{\lambda} Y_t. \tag{11.13}$$

The one period lead of (11.13) is

$$Y_t = -\frac{a}{\lambda} + \frac{1}{\lambda} Y_{t+1}. \tag{11.14}$$

Equation (11.14) can now be lead one period

$$Y_{t+1} = -\frac{a}{\lambda} + \frac{1}{\lambda} Y_{t+2}. \tag{11.15}$$

Substituting (11.15) into (11.14) gives

$$\begin{aligned} Y_t &= -\frac{a}{\lambda} + \frac{1}{\lambda}\left(-\frac{a}{\lambda} + \frac{1}{\lambda} Y_{t+2}\right) \\ &= -a\frac{1}{\lambda} - a\frac{1}{\lambda^2} + \frac{1}{\lambda^2} Y_{t+2}. \end{aligned} \tag{11.16}$$

The one period lead of (11.15) is

$$Y_{t+2} = -\frac{a}{\lambda} + \frac{1}{\lambda} Y_{t+3}. \tag{11.17}$$

Substituting (11.17) into (11.16) yields

$$\begin{aligned} Y_t &= -a\frac{1}{\lambda} - a\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\left(-\frac{a}{\lambda} + \frac{1}{\lambda} Y_{t+3}\right) \\ &= -a\frac{1}{\lambda} - a\frac{1}{\lambda^2} - a\frac{1}{\lambda^3} + \frac{1}{\lambda^3} Y_{t+3}. \end{aligned} \tag{11.18}$$

We can next substitute out Y_{t+3} and carry on successive forward substitutions to derive after $T-1$ iterations the following general formulation

$$Y_t = -a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^i + \left(\frac{1}{\lambda}\right)^T Y_{t+T} \tag{11.19}$$

Concentrating on the summation $S = -a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^i = -a\left[\left(\frac{1}{\lambda}\right) + \left(\frac{1}{\lambda}\right)^2 + \dots \left(\frac{1}{\lambda}\right)^T\right]$ for the moment, let's lead S by multiplying it by $\frac{1}{\lambda}$ and then subtract S from the product $S\frac{1}{\lambda}$

$$\begin{aligned}
S\left(\frac{1}{\lambda} - 1\right) &= \left(a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^i \frac{1}{\lambda}\right) - a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^i \\
&= a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^{i+1} - a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^i \\
&= a\sum_{i=2}^{T+1}\left(\frac{1}{\lambda}\right)^i - a\sum_{i=1}^{T}\left(\frac{1}{\lambda}\right)^i \qquad (11.20)\\
&= a\left[\left(\frac{1}{\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^3 + \dots \left(\frac{1}{\lambda}\right)^{T+1}\right] - a\left[\left(\frac{1}{\lambda}\right) + \left(\frac{1}{\lambda}\right)^2 + \dots \left(\frac{1}{\lambda}\right)^T\right] \\
&= a\left(\frac{1}{\lambda}\right)^{T+1} - a\left(\frac{1}{\lambda}\right).
\end{aligned}$$

Equation (11.20) can be written equivalently as

$$\begin{aligned}
S &= \frac{a\left(\frac{1}{\lambda}\right)^{T+1}}{\left(\frac{1}{\lambda} - 1\right)} - \frac{a\left(\frac{1}{\lambda}\right)}{\left(\frac{1}{\lambda} - 1\right)} \\
&= \frac{a\left(\frac{1}{\lambda}\right)^{T+1}}{\left(\frac{1-\lambda}{\lambda}\right)} - \frac{a\left(\frac{1}{\lambda}\right)}{\left(\frac{1-\lambda}{\lambda}\right)} \\
&= \frac{a\left(\frac{1}{\lambda}\right)^{T+1}\lambda}{1-\lambda} - \frac{a}{(1-\lambda)} \qquad (11.21)\\
&= \frac{a\left(\frac{1}{\lambda}\right)^{T}}{1-\lambda} - \frac{a}{(1-\lambda)} \\
&= \frac{a}{(1-\lambda)\lambda^T} - \frac{a}{(1-\lambda)}.
\end{aligned}$$

In equilibrium, $Y_t = Y_{t+1}$, accordingly our difference equation given by (11.14) can be rewritten as

$$\begin{aligned}
y &= -\frac{a}{\lambda} + \frac{1}{\lambda}y \\
&= \frac{-\frac{a}{\lambda}}{\frac{\lambda-1}{\lambda}} \qquad (11.22)\\
&= -\frac{a}{\lambda - 1} \\
&= \frac{a}{1-\lambda}.
\end{aligned}$$

Substituting (11.22) into (11.21) we can write S in equilibrium as

$$S = y\frac{1}{\lambda^T} - y. \tag{11.23}$$

Finally substituting (11.23) into (11.19) gives the general forward solution to the difference equation (11.1)

$$\begin{aligned} Y_t &= y - y\frac{1}{\lambda^T} + \left(\frac{1}{\lambda}\right)^T Y_{t+T} \\ &= y + \left(\frac{1}{\lambda}\right)^T (Y_{t+T} - y). \end{aligned} \tag{11.24}$$

As above the numeric solution to (11.14): $Y_t = -\frac{a}{\lambda} + \frac{1}{\lambda}Y_{t+1}$ requires that we know the terminal or final value Y_f for Y_{t+T}. If Y_f is exactly t periods after Y_t then $Y_t = y + \left(\frac{1}{\lambda}\right)^t (Y_f - y)$. Thus, we again have a general solution comprised of the sum of the *particular integral* and the *complementary function*, i.e. $y = \frac{a}{1-\lambda}$ and $\left(\frac{1}{\lambda}\right)^t (Y_f - y)$. If $0 < 1/\lambda < 1$ as $t \to \infty$, $Y_t \to y$ and the solution is stable or convergent.

12. Review 3: Solving Systems of Linear Difference Equations

12.1. Matrix notation

Consider the following two first order autonomous nonhomogeneous linear difference equations

$$\begin{aligned} x_t &= ax_{t-1} + by_{t-1} + e \\ y_t &= cx_{t-1} + dy_{t-1} + f. \end{aligned} \tag{12.1}$$

These equations can be rewritten in terms of matrices and vectors as

$$\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{g} \tag{12.2}$$

where,

$$\mathbf{z}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \ \mathbf{z}_{t-1} = \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix}, \ \mathbf{g} = \begin{bmatrix} e \\ f \end{bmatrix}.$$

12.2. Equilibrium

In equilibrium $x_t = x_{t-1} = x$ for all t and $y_t = y_{t-1} = y$ for all t. Accordingly

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}$$

or

$$\mathbf{z} = \mathbf{A}\mathbf{z} + \mathbf{g}. \tag{12.3}$$

An equilibrium solution exists for this system if

$$\begin{aligned} \mathbf{z} - \mathbf{A}\mathbf{z} &= \mathbf{g} \\ (\mathbf{I} - \mathbf{A})\mathbf{z} &= \mathbf{g} \\ \mathbf{z} &= (\mathbf{I} - \mathbf{A})^{-1}\mathbf{g}. \end{aligned}$$

Hence the equilibrium depends on the existence of $(\mathbf{I} - \mathbf{A})^{-1}$. In this example, the solution of the system in terms of x and y is

$$\begin{aligned} x &= \frac{e(1-d) + fb}{d(a - a/d - 1) - bc + 1} \\ y &= \frac{f(a-1) - ec}{d(a - a/d - 1) - bc + 1}. \end{aligned} \tag{12.4}$$

We can reduce the above system of non-homogeneous difference equations to a homogenous one by considering deviations from equilibrium. For example, we saw above that $\mathbf{z} = \mathbf{A}\mathbf{z} + \mathbf{g}$ was the equilibrium representation of the system $\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{g}$. If we define deviations from equilibrium as $\widehat{\mathbf{z}}_t = \mathbf{z}_t - \mathbf{z}$, then we can re-express the system as

$$\begin{aligned} \widehat{\mathbf{z}}_t &= \mathbf{z}_t - \mathbf{z} \\ &= (\mathbf{A}\mathbf{z}_{t-1} + \mathbf{g}) - (\mathbf{A}\mathbf{z} + \mathbf{g}) \\ &= \mathbf{A}(\mathbf{z}_{t-1} - \mathbf{z}) \\ &= \mathbf{A}\widehat{\mathbf{z}}_{t-1} \end{aligned} \tag{12.5}$$

where,

$$\begin{aligned} \widehat{\mathbf{z}}_t &= \begin{bmatrix} \widehat{x}_t \\ \widehat{y}_t \end{bmatrix} \\ &= \begin{bmatrix} x_t - x \\ y_t - y \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \widehat{\mathbf{z}}_{t-1} &= \begin{bmatrix} \widehat{x}_{t-1} \\ \widehat{y}_{t-1} \end{bmatrix} \\ &= \begin{bmatrix} x_{t-1} - x \\ y_{t-1} - y \end{bmatrix}. \end{aligned}$$

In equilibrium $\widehat{x}_t = \widehat{x}_{t-1} = \widehat{x}$ for all t and $\widehat{y}_t = \widehat{y}_{t-1} = \widehat{y}$ for all t. Accordingly

$$\begin{aligned} \widehat{\mathbf{z}} - \mathbf{A}\widehat{\mathbf{z}} &= 0 \\ (\mathbf{I} - \mathbf{A})\widehat{\mathbf{z}} &= 0 \\ \widehat{\mathbf{z}} &= (\mathbf{I} - \mathbf{A})^{-1}0 \\ &= 0. \end{aligned} \tag{12.6}$$

Hence the solution to the autonomous system requires that $\widehat{x}$=0 and $\widehat{y}$=0. It can easily be verified that the only values of $\widehat{x}$ and $\widehat{y}$ which satisfy the above system are indeed zero.

12.3. Stability of discrete systems

Given that we have established the existence of an equilibrium for the above non-homogenous and homogenous systems we next need to solve the respective systems to set out the conditions required for dynamically stable. The solutions to both of these systems is straightforward and can be carried out by repeated substitution as shown above.

12.3.1. Non-homogenous system

For example in the non-homogenous case, the general solution can be found as follows

$$\begin{aligned}
\mathbf{z}_t &= \mathbf{A}\mathbf{z}_{t-1} + \mathbf{g} \\
&= \mathbf{A}(\mathbf{A}\mathbf{z}_{t-2} + \mathbf{g}) + \mathbf{g} = \mathbf{A}^2\mathbf{z}_{t-2} + \mathbf{A}\mathbf{g} + \mathbf{g} \\
&= \mathbf{A}^2(\mathbf{A}\mathbf{z}_{t-3}+\mathbf{g}) + \mathbf{A}\mathbf{g} + \mathbf{g} = \mathbf{A}^3\mathbf{z}_{t-3}+\mathbf{A}^2\mathbf{g} + \mathbf{A}\mathbf{g} + \mathbf{g} \qquad (12.7)\\
&. \\
&. \\
&= \mathbf{A}^t\mathbf{z}_0+(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{t-1})\mathbf{g}.
\end{aligned}$$

where $\mathbf{z}_0$ is the initial values of the vector $\mathbf{z}$.

12.3.2. Homogenous system

In the homogenous case the general solution is given by

$$\begin{aligned}
\widehat{\mathbf{z}}_t &= \mathbf{A}\widehat{\mathbf{z}}_{t-1} \\
&= \mathbf{A}(\mathbf{A}\widehat{\mathbf{z}}_{t-2}) = \mathbf{A}^2\widehat{\mathbf{z}}_{t-2} \\
&= \mathbf{A}^2(\mathbf{A}\widehat{\mathbf{z}}_{t-3}) = \mathbf{A}^3\widehat{\mathbf{z}}_{t-3} \qquad (12.8)\\
&. \\
&. \\
&= \mathbf{A}^t\widehat{\mathbf{z}}_0
\end{aligned}$$

where $\widehat{\mathbf{z}}_0$ is the initial values of the vector $\widehat{\mathbf{z}}$.

Note that, if $b = c = 0$, then $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is diagonal and (12.5) reduces to two independent homogenous equations in a single variable

$$\begin{aligned}
\widehat{x}_t &= a\widehat{x}_{t-1} \\
\widehat{y}_t &= d\widehat{y}_{t-1}
\end{aligned}$$

with general solutions readily given by

$$\begin{aligned}\widehat{x}_t &= a^t \widehat{x}_0 \\ \widehat{y}_t &= d^t \widehat{y}_0.\end{aligned}$$

Coefficient matrices in typical economic systems are not diagonal, however we will see below that application of a diagonalizing transformation will allow us to rewrite many systems in this form. This permits us to easily solve the transformed system and then to recover the solution to the pre-transformed system by inverting the diagonalizing transformation. We will see below that to diagonalize $\mathbf{A}$ we need to find an invertible matrix $\mathbf{V}$ such that $\mathbf{D} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$ where $\mathbf{V}$ is composed of the eigenvectors associated with the eigenvalues of $\mathbf{A}$, i.e. $\mathbf{D}$.

12.3.3. Eigenvalues and eigenvectors

Given that linear non-homogeneous systems can be re-expressed as homogeneous ones, without loss of generality, we will only concentrate on the latter in our subsequent discussion regarding stability. It will be clear by now that the matrix $\mathbf{A}$ is very important when assessing the stability of linear systems of difference equations. Under appropriate conditions $\mathbf{A}$ can be diagonalized or decomposed into another matrix known as the Jordan canonical form, which is expressed in terms of the eigenvalues and eigenvectors of $\mathbf{A}$.

Let's start by reviewing the concepts of eigenvalues and eigenvectors. Consider the following general linear system

$$\mathbf{w} = \mathbf{A}\mathbf{x}. \tag{12.9}$$

The matrix $\mathbf{A}$ acts to transform the vector $\mathbf{x}$ to the vector $\mathbf{w}$. Now let's next suppose that $\mathbf{x}$ is transformed into a multiple of itself, e.g. $\mathbf{w} = \lambda\mathbf{x}$ where λ is a scalar of proportionality. Then it follows that

$$\begin{aligned}\mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \qquad (12.10) \\ (\mathbf{A}-\lambda\mathbf{I})\mathbf{x} &= 0.\end{aligned}$$

The solution to (12.10) will only be non-zero if λ is chosen such that

$$\varphi(\lambda) = \det(\mathbf{A}-\lambda\mathbf{I}) = 0. \tag{12.11}$$

The values of λ which satisfy (12.11) are called the eigenvalues or characteristic values or characteristic roots of $\mathbf{A}$ and the solutions to the system (i.e. for $\mathbf{x}$) that are obtained using these values of λ are called the eigenvectors corresponding to the eigenvalues.

For example, the eigenvalues of **A** (using the 2x2 example from 12.5 above) are found as follows

$$\begin{aligned}\det(\mathbf{A}-\lambda\mathbf{I}) &= \left|\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right| = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0 \\ &= \lambda^2 - (a+d)\lambda + (ad-bc) = 0 \\ &= \lambda^2 - tr(\mathbf{A})\lambda + \det(\mathbf{A}) = 0. \end{aligned} \tag{12.12}$$

The solution to the polynomial or characteristic equation can be obtained by application of the quadratic formula, e.g.

$$\begin{aligned}\lambda_i &= \frac{tr(\mathbf{A}) \pm \sqrt{tr(\mathbf{A})^2 - 4\det(\mathbf{A})}}{2} \\ &= \frac{1}{2}(a+d) \pm \sqrt{a^2 - 2ad + d^2 + 4bc}\end{aligned}$$

where the subscript $i = r, s$ refers to the two eigenvalues of the characteristic equation. If $tr(\mathbf{A})^2 > 4\det(\mathbf{A})$ then the roots are real and distinct; $tr(\mathbf{A})^2 = 4\det(\mathbf{A})$ the roots are real and equal; and if $tr(\mathbf{A})^2 < 4\det(\mathbf{A})$ then the roots are complex conjugate.

The eigenvectors corresponding to these eigenvalues are determined by substituting for each eigenvalue in the following equation[17] and solving for $\mathbf{v}^i$

$$(\mathbf{A}-\lambda_i\mathbf{I})\mathbf{v}^i = 0. \tag{12.13}$$

For $\lambda = r$ we have

$$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \left(\frac{1}{2}(a+d) + \sqrt{a^2 - 2ad + d^2 + 4bc}\right)\mathbf{I}\right)\begin{bmatrix} v_1^r \\ v_2^r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

hence

$$\mathbf{v}^r = \begin{bmatrix} v_1^r \\ v_2^r \end{bmatrix} = \begin{bmatrix} \frac{\frac{1}{2}(a+d)+\frac{1}{2}\sqrt{a^2-2ad+d^2+4bc}}{c} \\ 1 \end{bmatrix}.$$

Likewise for $\lambda = s$ we have

$$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \left(\frac{1}{2}(a+d) - \sqrt{a^2 - 2ad + d^2 + 4bc}\right)\mathbf{I}\right)\begin{bmatrix} v_1^s \\ v_2^s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

hence

$$\mathbf{v}^s = \begin{bmatrix} v_1^s \\ v_2^s \end{bmatrix} = \begin{bmatrix} \frac{\frac{1}{2}(a+d)-\frac{1}{2}\sqrt{a^2-2ad+d^2+4bc}}{c} \\ 1 \end{bmatrix}.$$

[17] Note that this equation takes the same form as (12.10).

12.3.4. Transformation of the A matrix

Given the above background concerning eigenvalues and eigenvectors we are now in a position to address the issue of stability using our linear homogenous system $\widehat{\mathbf{z}}_t = \mathbf{A}\widehat{\mathbf{z}}_{t-1}$ with the general solution $\widehat{\mathbf{z}}_t$=$\mathbf{A}^t\widehat{\mathbf{z}}_0$. First let's be more explicit about the diagonalization or the Jordon canonical decomposition referred to above.

Theorem (2-variable case)

If the eigenvalues of the matrix $\mathbf{A}$ are r and s obtained from $\det(\mathbf{A}-\lambda\mathbf{I}) = 0$ such that $r \neq s$ then there exists a matrix $\mathbf{V}$=[$\mathbf{v}^r$,$\mathbf{v}^s$] composed of the eigenvectors associated with r and s respectively, such that

$$\mathbf{D} = \begin{bmatrix} r & 0 \\ 0 & s \end{bmatrix} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}, \tag{12.14}$$

where $\mathbf{v}^r = \begin{bmatrix} v_1^r \\ v_2^r \end{bmatrix}$ and $\mathbf{v}^s = \begin{bmatrix} v_1^s \\ v_2^s \end{bmatrix}$.

According we can see that $\mathbf{A}$ can be decomposed into Jordan canonical form which is comprised of its eigenvalues and eigenvectors, e.g.

$$\begin{aligned} \mathbf{D} &= \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \\ \mathbf{V}\mathbf{D}\mathbf{V}^{-1} &= \mathbf{V}(\mathbf{V}^{-1}\mathbf{A}\mathbf{V})\mathbf{V}^{-1} \\ &= \mathbf{A}. \end{aligned}$$

Moreover it follows that

$$\begin{aligned} \mathbf{A}^2 &= \mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{V}\mathbf{D}\mathbf{V}^{-1} = \mathbf{V}\mathbf{D}^2\mathbf{V}^{-1} \\ \mathbf{A}^3 &= \mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{V}\mathbf{D}^2\mathbf{V}^{-1} = \mathbf{V}\mathbf{D}\mathbf{D}^2\mathbf{V}^{-1} = \mathbf{V}\mathbf{D}^3\mathbf{V}^{-1} \\ &\cdot \\ &\cdot \\ \mathbf{A}^t &= \mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{V}\mathbf{D}^{t-1}\mathbf{V}^{-1} = \mathbf{V}\mathbf{D}\mathbf{D}^{t-1}\mathbf{V}^{-1} = \mathbf{V}\mathbf{D}^t\mathbf{V}^{-1}. \end{aligned}$$

Therefore our general solution $\widehat{\mathbf{z}}_t = \mathbf{A}^t\widehat{\mathbf{z}}_0$ can be reexpressed as

$$\begin{aligned} \widehat{\mathbf{z}}_t &= \mathbf{V}\mathbf{D}^t\mathbf{V}^{-1}\widehat{\mathbf{z}}_0 \\ &= \mathbf{V}\begin{bmatrix} r^t & 0 \\ 0 & s^t \end{bmatrix}\mathbf{V}^{-1}\widehat{\mathbf{z}}_0. \end{aligned} \tag{12.15}$$

We can simplify $\widehat{\mathbf{z}}_t = \mathbf{A}^t\widehat{\mathbf{z}}_0$ even further by pre-multiplying both sides by $\mathbf{V}^{-1}$

$$\begin{aligned}
\mathbf{V}^{-1}\widehat{\mathbf{z}}_t &= \mathbf{V}^{-1}\mathbf{A}^t\widehat{\mathbf{z}}_0 \\
&= \mathbf{V}^{-1}\mathbf{A}^t\mathbf{V}\mathbf{V}^{-1}\widehat{\mathbf{z}}_0, \text{ where } \mathbf{V}\mathbf{V}^{-1} = \mathbf{I} \\
\widetilde{\widehat{\mathbf{z}}}_t &= \mathbf{V}^{-1}\mathbf{A}^t\mathbf{V}\widetilde{\widehat{\mathbf{z}}}_0, \text{ where } \widetilde{\widehat{\mathbf{z}}}_t = \mathbf{V}^{-1}\widehat{\mathbf{z}}_t \text{ and } \widetilde{\widehat{\mathbf{z}}}_0 = \mathbf{V}^{-1}\widehat{\mathbf{z}}_0 \\
&= \mathbf{D}^t\widetilde{\widehat{\mathbf{z}}}_0 \\
&= \begin{bmatrix} r^t & 0 \\ 0 & s^t \end{bmatrix}\widetilde{\widehat{\mathbf{z}}}_0
\end{aligned}$$

or equivalently

$$\begin{aligned}
\widetilde{\widehat{x}}_t &= r^t\widetilde{\widehat{x}}_0 \\
\widetilde{\widehat{y}}_t &= s^t\widetilde{\widehat{y}}_0.
\end{aligned}$$

Since $\widetilde{\widehat{\mathbf{z}}}_t = \mathbf{V}^{-1}\widehat{\mathbf{z}}_t \Longrightarrow \widehat{\mathbf{z}}_t = \mathbf{V}\widetilde{\widehat{\mathbf{z}}}_t$, we can recover the solution to the original model as follows

$$\mathbf{V}\widetilde{\widehat{\mathbf{z}}}_t = \mathbf{V}\begin{bmatrix} r^t & 0 \\ 0 & s^t \end{bmatrix}\widetilde{\widehat{\mathbf{z}}}_0$$

or

$$\begin{bmatrix} \widehat{x}_t \\ \widehat{y}_t \end{bmatrix} = \begin{bmatrix} v_1^r & v_1^s \\ v_2^r & v_2^s \end{bmatrix}\begin{bmatrix} r^t\widetilde{\widehat{x}}_0 \\ s^t\widetilde{\widehat{y}}_0 \end{bmatrix}$$

or equivalently

$$\begin{aligned}
\widehat{x}_t &= v_1^r r^t\widetilde{\widehat{x}}_0 + v_1^s s^t\widetilde{\widehat{y}}_0 \\
\widehat{y}_t &= v_2^r r^t\widetilde{\widehat{x}}_0 + v_2^s s^t\widetilde{\widehat{y}}_0.
\end{aligned}$$

To denote that the starting values or boundary conditions can take on any constant value we can redefine $\widetilde{\widehat{x}}_0$ and $\widetilde{\widehat{y}}_0$ as c_1 and c_2 respectively,

$$\begin{aligned}
\widehat{x}_t &= c_1v_1^r r^t + c_2v_1^s s^t \\
\widehat{y}_t &= c_1v_2^r r^t + c_2v_2^s s^t.
\end{aligned} \tag{12.16}$$

12.3.5. Stability conditions for two-variable systems

- Assuming that r and s are real and distinct, we can readily see that for any value of c_1 and c_2, if $|r| < 1$ and $|s| < 1$ then as $t \to \infty$, $(r^t, s^t \to 0)$ hence $\widehat{x}_t \to 0$. The same applies to $\widehat{y}_t$. Consequently the system tends to the fixed or equilibrium point of zero established above and is dynamically stable.

- If $|r| > 1$ and $|s| > 1$ then $(\widehat{x}_t, \widehat{y}_t \to \pm\infty)$ depending on the signs of r and s and is dynamically unstable (unless $c_1 = c_2 = 0$, i.e. the system starts in the steady-state).

- If $|r| > 1$ and $|s| < 1$ then the system is generally (i.e. for most boundary conditions) dynamically unstable since the system will tend to be dominated by the root which is greater than unity and will according tend to plus or minus infinity depending on its sign. However there is one non-stationary solution known as the *saddle path* which converges to the steady-state. To obtain it, we impose a boundary condition to eliminate the impact of the explosive root, i.e. here we would set $c_1 = 0$. Given this restriction our general solution given by (12.16) becomes

$$\begin{aligned}\widehat{x}_t &= c_2 v_1^s s^t \\ \widehat{y}_t &= c_2 v_2^s s^t.\end{aligned} \tag{12.17}$$

Eliminating $c_2 s^t$ from (12.17) gives the equation of the saddle path

$$\widehat{x}_t = \frac{v_1^s}{v_2^s}\widehat{y}_t.$$

Saddle path solution are an important feature of rational expectations theory. In a two variable context where say y is predetermined and x is not, there will only be one unique initial value of x associated with a given initial value y for which the system will converge to the steady-state. Since x is non-predetermined, under rational expectations, it is free to "jump" precisely to the value needed to put the system on the stable saddle path.

13. Review 4: State Space Models and the Kalman Filter

13.1. State space setup

Consider the general state space model[18]

$$\boldsymbol{\alpha}_t = \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t \boldsymbol{\eta}_t; \tag{13.1.1}$$

$$\mathbf{y}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t; \quad t = 1, ..., N \tag{13.1.2}$$

where $\boldsymbol{\alpha}_t$ is the unobservable $m \times 1$ state vector at time t, $\mathbf{y}_t$ is the $n \times 1$ measurement vector of observations, $\boldsymbol{\eta}_t$ and $\boldsymbol{\varepsilon}_t$ are $g \times 1$ and $n \times 1$ vectors of serially uncorrelated Gaussian disturbances with mean $\mathbf{0}$ and covariance matrices $\mathbf{Q}_t$ and $\mathbf{H}_t$.[19] The *system matrices* are comprised of the $m \times m$ *transition matrix* $\mathbf{T}_t$, the $m \times g$ matrix $\mathbf{R}_t$, the $n \times m$ matrix $\mathbf{Z}_t$, the $n \times 1$ and $m \times 1$ non-stochastic vectors $\mathbf{c}_t$ and $\mathbf{d}_t$, and the covariance matrices $\mathbf{Q}_t$ and $\mathbf{H}_t$. Note that the matrices $\mathbf{T}_t$, $\mathbf{c}_t$, $\mathbf{R}_t$, $\mathbf{Z}_t$ and $\mathbf{d}_t$ are non-stochastic.

13.2. Elements of the system matrices

The transition and measurement equation systems in (13.1.1-13.1.2) are given by the economic model under analysis with the former determining the dynamics of the $m \times 1$ state vector, $\boldsymbol{\alpha}_t$. Accordingly the $\mathbf{T}_t$ and $\mathbf{Z}_t$ matrices are comprised of $0's$, $1's$ and the structural economic parameters of interest. If these parameters and those in the $\mathbf{c}_t$ vector do not vary over time then $\mathbf{T}_t$, $\mathbf{Z}_t$ and $\mathbf{c}_t$ are constant for $t = 1, ..., N$. The $\mathbf{R_t}$ matrix controls the number of shocks entering the state vector and contains $0's$ and $1's$. If we do not allow for time-varying parameters on the non-stochastic variables (e.g. trends, dummy variables, etc.) in $\mathbf{d}_t$ then it will also be constant for all t. Finally, if the vectors of error variances in the state and measurement equations are time invariant then the covariance matrices $\mathbf{Q}_t$ and $\mathbf{H}_t$ will be constant over time. The optimally estimated elements of $\mathbf{T}_t$, $\mathbf{c}_t$, $\mathbf{d}_t$, $\mathbf{Q}_t$ and $\mathbf{H}_t$ are known as the *hyperparameters*.

[18] The following treatment is based on Harvey (1992, pp. 100-130), Lütkepohl (1991, pp. 418-419, 428-435) and Woitek (2004, pp. 55-59).

[19] More general formulations could easily allow for scalar and vector AR and MA error processes in the measurement and transitions equations.

13.3. Kalman filter

Consider the conditional distribution of the state vector $\boldsymbol{\alpha}_t$ given information at time s. The mean, and covariance matrix, of this distribution can be defined as

$$\mathbf{a}_{t|s} \equiv E_s(\boldsymbol{\alpha}_t), \tag{13.2.1}$$

$$\mathbf{P}_{t|s} \equiv E_s\left(\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|s}\right)\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|s}\right)'\right]\right) \tag{13.2.2}$$

where the expectations operator indicates that expectations are formed using the conditional distribution for that period.

To obtain the one-step ahead mean, $\mathbf{a}_{t|t-1}$ and covariance, $\mathbf{P}_{t|t-1}$ of $\boldsymbol{\alpha}_t$ we use the conditional distribution implied by setting $s = t-1$. This yields the following *prediction equations*

$$\begin{aligned}\mathbf{a}_{t|t-1} &\equiv E_{t-1}(\boldsymbol{\alpha}_t) \\ &= \mathbf{T}_t\mathbf{a}_{t-1} + \mathbf{c}_t,\end{aligned} \tag{13.3.1}$$

$$\begin{aligned}\mathbf{P}_{t|t-1} &\equiv E_{t-1}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)'\right] \\ &= \mathbf{T}_t\mathbf{P}_{t-1}\mathbf{T}_t' + \mathbf{R}_t\mathbf{Q}_t\mathbf{R}_t'.\end{aligned} \tag{13.3.2}$$

Derivations of (13.3.1) and (13.3.2) are given below.

To calculate the above *prediction equations* we need to assume initial values for the elements of the state vector in the previous period, $\mathbf{a}_{t-1}$ and for the system matrices $\mathbf{T}_t$, $\mathbf{c}_t$, $\mathbf{R}_t$, and $\mathbf{Q}_t$. The second step in calculating the Kalman Filter is to revise the estimation error from step-one using the *updating equations* which are based on actual observations of y available at time t. The updating equations are given by

$\mathbf{y}_{t\|t-1}$	$= \mathbf{Z}_t\mathbf{a}_{t\|t-1} + \mathbf{d}_t;$	*1-step ahead estimate of* $\mathbf{y}_t$	(13.4.1)
$\mathbf{F}_t$	$= \mathbf{Z}_t\mathbf{P}_{t\|t-1}\mathbf{Z}_t' + \mathbf{H}_t;$	*estimated cov matrix of* $\boldsymbol{\nu}_t$	(13.4.2)
$\boldsymbol{\nu}_t$	$= \mathbf{y}_t - \mathbf{y}_{t\|t-1};$	*observation vector estimation error*	(13.4.3)
$\mathbf{a}_t$	$= \mathbf{a}_{t\|t-1} + \mathbf{K}_t\boldsymbol{\nu}_t;$	*updating of the state vector*	(13.4.4)
$\mathbf{K}_t$	$= \mathbf{P}_{t\|t-1}\mathbf{Z}_t'\mathbf{F}_t^{-1};$	*Kalman gain*	(13.4.5)
$\mathbf{P}_t$	$= \mathbf{P}_{t\|t-1} - \mathbf{K}_t\mathbf{Z}_t\mathbf{P}_{t\|t-1};$	*updating of state-cov matrix*	(13.4.6)

Further details motivating $\boldsymbol{\nu}_t$ and $\mathbf{a}_t$ as well as derivations for $\mathbf{y}_{t|t-1}$, $\mathbf{F}_t$, $\mathbf{K}_t$ and $\mathbf{P}_t$ are given below. To undertake the calculations required in the *updating equations* we need to assume initial values for the elements of $\mathbf{Z}_t$, $\mathbf{d}_t$ and $\mathbf{H}_t$.

13.3.1. Derivation of the conditional state mean $\mathbf{a}_{t|t-1}$

The one-step ahead mean, $\mathbf{a}_{t|t-1}$ defined in (13.3.1) can be derived as follows

$$
\begin{aligned}
\mathbf{a}_{t|t-1} &\equiv E_{t-1}(\boldsymbol{\alpha}_t) \\
&= E_{t-1}(\mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t\boldsymbol{\eta}_t) \\
&= \mathbf{T}_t E_{t-1}(\boldsymbol{\alpha}_{t-1}) + E_{t-1}\mathbf{c}_t + \mathbf{R}_t E_{t-1}(\boldsymbol{\eta}_t) \\
&= \mathbf{T}_t\mathbf{a}_{t-1} + \mathbf{c}_t.
\end{aligned}
$$

13.3.2. Derivation of the conditional state covariance matrix $\mathbf{P}_{t|t-1}$

The one-step ahead covariance matrix, $\mathbf{P}_{t|t-1}$ defined in (13.3.2) can be derived as follows

$$
\begin{aligned}
\mathbf{P}_{t|t-1} &\equiv E_{t-1}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)'\right] \\
&= E_{t-1}\left[(\mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t\boldsymbol{\eta}_t - \mathbf{T}_t\mathbf{a}_{t-1} - \mathbf{c}_t)(\mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t\boldsymbol{\eta}_t - \mathbf{T}_t\mathbf{a}_{t-1} - \mathbf{c}_t)'\right] \\
&= E_{t-1}\left[(\mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{R}_t\boldsymbol{\eta}_t - \mathbf{T}_t\mathbf{a}_{t-1})(\mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{R}_t\boldsymbol{\eta}_t - \mathbf{T}_t\mathbf{a}_{t-1})'\right] \\
&= E_{t-1}\left\{[\mathbf{T}_t(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1}) + \mathbf{R}_t\boldsymbol{\eta}_t][\{\mathbf{T}_t(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})\}' + (\mathbf{R}_t\boldsymbol{\eta}_t)']\right\} \\
&= E_{t-1}\left\{[\mathbf{T}_t(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1}) + \mathbf{R}_t\boldsymbol{\eta}_t][(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})'\mathbf{T}_t' + \boldsymbol{\eta}_t'\mathbf{R}_t']\right\} \\
&= \mathbf{T}_t E_{t-1}\left[(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})'\right]\mathbf{T}_t' + \mathbf{T}_t E_{t-1}\left[(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})\boldsymbol{\eta}_t'\right]\mathbf{R}_t' + \\
&\quad \mathbf{R}_t E_{t-1}\left(\boldsymbol{\eta}_t(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})'\right)\mathbf{T}_t' + \mathbf{R}_t E_{t-1}(\boldsymbol{\eta}_t\boldsymbol{\eta}_t')\mathbf{R}_t' \\
&= \mathbf{T}_t \underbrace{E_{t-1}\left[(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})'\right]}_{=\mathbf{P}_{t-1}}\mathbf{T}_t' + \mathbf{R}_t \underbrace{E_{t-1}[\boldsymbol{\eta}_t\boldsymbol{\eta}_t']}_{=\mathbf{Q}_t}\mathbf{R}_t' \\
&= \mathbf{T}_t\mathbf{P}_{t-1}\mathbf{T}_t' + \mathbf{R}_t\mathbf{Q}_t\mathbf{R}_t'.
\end{aligned}
$$

Note that the two terms involving the error, $\boldsymbol{\eta}_t$, i.e. $\mathbf{T}_t E_{t-1}\left[(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})\boldsymbol{\eta}_t'\right]\mathbf{R}_t'$ and $\mathbf{R}_t E_{t-1}\left(\boldsymbol{\eta}_t(\boldsymbol{\alpha}_{t-1} - \mathbf{a}_{t-1})'\right)\mathbf{T}_t'$, are equal to zero since the expected value of the product of two independent random variables is equal to the expected value of the first variable times the expected value of the second and $E_{t-1}[\boldsymbol{\eta}_t] = 0$.

13.3.3. Derivation of the conditional measurement mean $\mathbf{y}_{t|t-1}$

The one-step ahead measurement mean, $\mathbf{y}_{t|t-1}$ defined in (13.4.1) can be derived as follows

$$\begin{aligned}\mathbf{y}_{t|t-1} &\equiv E_{t-1}(\mathbf{y}_t) \\ &= E_{t-1}(\mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t) \\ &= \mathbf{Z}_t(E_{t-1}\boldsymbol{\alpha}_t) + E_{t-1}\mathbf{d}_t + \mathbf{E_{t-1}}\boldsymbol{\varepsilon}_t \\ &= \mathbf{Z}_t\mathbf{a}_{t|t-1} + \mathbf{d}_t.\end{aligned}$$

13.3.4. Derivation of the conditional observation vector error covariance matrix, $\mathbf{F}_t$

The error covariance matrix, $\mathbf{F}_t$ defined in (13.4.2) can be derived as follows

$$\begin{aligned}\mathbf{F}_t &\equiv E_{t-1}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)'\right] \\ &= E_{t-1}\left[\left(\mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t - \mathbf{Z}_t\mathbf{a}_{t|t-1} - \mathbf{d}_t\right)\left(\mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t - \mathbf{Z}_t\mathbf{a}_{t|t-1} - \mathbf{d}_t\right)'\right] \\ &= E_{t-1}\left[\left(\mathbf{Z}_t(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\varepsilon}_t\right)\left(\mathbf{Z}_t(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\varepsilon}_t\right)'\right] \\ &= E_{t-1}\left[\left(\mathbf{Z}_t(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\varepsilon}_t\right)\left((\mathbf{Z}_t\boldsymbol{\alpha}_t)' - \left(\mathbf{Z}_t\mathbf{a}_{t|t-1}\right)' + \boldsymbol{\varepsilon}_t'\right)\right] \\ &= E_{t-1}\left[\left(\mathbf{Z}_t(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\varepsilon}_t\right)\left(\boldsymbol{\alpha}_t{}'\mathbf{Z}_t' - \mathbf{a}_{t|t-1}'\mathbf{Z}_t' + \boldsymbol{\varepsilon}_t'\right)\right] \\ &= E_{t-1}\left[\left(\mathbf{Z}_t(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\varepsilon}_t\right)\left((\boldsymbol{\alpha}_t' - \mathbf{a}_{t|t-1}')\mathbf{Z}_t' + \boldsymbol{\varepsilon}_t'\right)\right] \\ &= \mathbf{Z}_t E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t' - \mathbf{a}_{t|t-1}')\right]\mathbf{Z}_t' + \mathbf{Z}_t E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\boldsymbol{\varepsilon}_t'\right] + \\ &\quad E_{t-1}\left[\boldsymbol{\varepsilon}_t(\boldsymbol{\alpha}_t' - \mathbf{a}_{t|t-1}')\mathbf{Z}_t'\right] + E_{t-1}\left[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right] \\ &= \mathbf{Z}_t \underbrace{E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t' - \mathbf{a}_{t|t-1}')\right]}_{\mathbf{P}_{t|t-1}}\mathbf{Z}_t' + \underbrace{E_{t-1}\left[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right]}_{\mathbf{H}_t} \\ \mathbf{F}_t &= \mathbf{Z}_t\mathbf{P}_{t|t-1}\mathbf{Z}_t' + \mathbf{H}_t.\end{aligned}$$

13.3.5. Intuition underlying the Kalman filter

The Kalman filter provides the best linear approximation of the true state vector $\boldsymbol{\alpha}_t$, if the state vector estimation error, $(\boldsymbol{\alpha}_t - \mathbf{a}_t)$ is independent of the past and present observations $\mathbf{y}_t$, i.e.

$$\mathrm{Cov}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_t\right), \mathbf{y}_s\right] = \mathbf{0};\ s = 1, \ldots, t. \tag{13.5}$$

The Kalman gain, $\mathbf{K}_t$ defined in (13.4.5) is derived to ensure that the above condition holds. To understand how this works, we start by assuming that the state vector estimation error, $(\boldsymbol{\alpha}_t - \mathbf{a}_t)$ is equal to the difference between the true state vector, $\boldsymbol{\alpha}_t$ and the prediction of the state vector based on information in the previous period, $\mathbf{a}_{t|t-1}$ *net* of a proportion, $\mathbf{K}_t$ of the observation vector estimation error, $\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)$, i.e.

$$(\boldsymbol{\alpha}_t - \mathbf{a}_t) = (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right). \tag{13.6}$$

Equation (13.6) implies the state updating equation given by equation (13.4.4), i.e.

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{K}_t\boldsymbol{\nu}_t$$

where $\boldsymbol{\nu}_t = \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)$, was defined in 13.4.3.

The above discussion also implies that for the observations $\mathbf{y}_s$, $s = 1, \ldots, t-1$ and any arbitrary matrix $\mathbf{K}_t$ the following condition must hold

$$\begin{aligned} \text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_t), \mathbf{y}_s\right] &= \text{Cov}\left[\left\{(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right\}, \mathbf{y}_s\right] = \mathbf{0} \\ &\implies \text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), \mathbf{y}_s\right] - \mathbf{K}_t\text{Cov}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right), \mathbf{y}_s\right] = \mathbf{0} \\ s &= 1, \ldots, t-1. \end{aligned} \tag{13.7}$$

13.3.6. Derivation of the Kalman gain, $\mathbf{K}_t$

To derive $\mathbf{K}_t$, defined in (13.4.5) we require the condition given in (13.7) plus the final condition that the state vector estimation error, $(\boldsymbol{\alpha}_t - \mathbf{a}_t)$ does not depend on the observation vector estimation error $\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)$, i.e.

$$\text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_t), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] = \mathbf{0}.$$

Substituting for $(\boldsymbol{\alpha}_t - \mathbf{a}_t)$ using (13.6) the above condition becomes

$$\text{Cov}\left[\left\{(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right\}, \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] = \mathbf{0},$$

and implies

$$\begin{aligned} &\text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] - \\ &\qquad \mathbf{K}_t\underbrace{\text{Cov}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]}_{\mathbf{F}_t} = \mathbf{0}. \end{aligned} \tag{13.8}$$

From subsection 13.3.4 we know $\text{Cov}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] \equiv \mathbf{F}_t$. Moreover it can be show that the first term in (13.8) can be written as

$$\text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] = \text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right]\mathbf{Z}_t'.$$

This result can be obtained by making use of (13.4.1): $\mathbf{y}_{t|t-1} = \mathbf{Z}_t \mathbf{a}_{t|t-1} + \mathbf{d}_t$ and (13.1.2): $\mathbf{y}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t$ as follows:

$$\begin{aligned}
& Cov\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] \\
\equiv\ & E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)'\right] \\
=\ & E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) \left(\mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t - \mathbf{Z}_t \mathbf{a}_{t|t-1} - \mathbf{d}_t\right)'\right] \\
=\ & E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) \left(\boldsymbol{\alpha}_t' \mathbf{Z}_t' - \mathbf{a}_{t|t-1}' \mathbf{Z}_t' + \boldsymbol{\varepsilon}_t'\right)\right] \\
=\ & E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) \left(\left(\boldsymbol{\alpha}_t' - \mathbf{a}_{t|t-1}'\right) \mathbf{Z}_t' + \boldsymbol{\varepsilon}_t'\right)\right] \\
=\ & E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) \left(\boldsymbol{\alpha}_t' - \mathbf{a}_{t|t-1}'\right)\right] \mathbf{Z}_t' + E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\boldsymbol{\varepsilon}_t'\right] \\
=\ & \text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] \mathbf{Z}_t' \quad \mathit{QED.}
\end{aligned}$$

Substituting this result and $\mathbf{F}_t$ into (13.8) gives

$$\text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] \mathbf{Z}_t' - \mathbf{K}_t \mathbf{F}_t = \mathbf{0}.$$

Moreover from subsection 13.3.2 we know $\text{Cov}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] \equiv \mathbf{P}_{t|t-1}$. Accordingly the above expression can be further simplified to give the Kalman gain

$$\begin{aligned}
\mathbf{P}_{t|t-1}\mathbf{Z}_t' - \mathbf{K}_t \mathbf{F}_t &= 0 \\
\mathbf{K}_t \mathbf{F}_t &= \mathbf{P}_{t|t-1}\mathbf{Z}_t' \\
\mathbf{K}_t &= \mathbf{P}_{t|t-1}\mathbf{Z}_t'(\mathbf{F}_t)^{-1}.
\end{aligned}$$

as defined in (13.4.5).

13.3.7. Derivation of the updating covariance matrix

The updating formula for $\mathbf{P}_t$ defined in (13.4.6) can be obtained as follows

$$\begin{aligned}
\mathbf{P}_t &= \text{Var}\left[\boldsymbol{\alpha}_t - \mathbf{a}_t\right] \\
&= \text{Var}\left[\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1} - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] \\
&\equiv E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]^2 \\
&= E_{t-1}\left\{\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) - \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]'\right\} \\
&= E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})'\right] - \mathbf{K}_t E_{t-1}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})'\right] - \\
&\quad E_{t-1}\left\{(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\left[\mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]'\right\} + E_{t-1}\left\{\mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\left[\mathbf{K}_t\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]'\right\} \\
&= E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})'\right] - \mathbf{K}_t E_{t-1}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})'\right] - \\
&\quad E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)'\right]\mathbf{K}_t' + \mathbf{K}_t E_{t-1}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)'\right]\mathbf{K}_t'
\end{aligned}$$

Given $E_{t-1}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)'\right] \equiv \mathrm{Cov}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right]$;
$E_{t-1}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)'\right] \equiv \mathrm{Var}\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)$ and from (13.8)
$\mathrm{Cov}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right), \left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\right] = \mathbf{K}_t \mathrm{Var}\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)$, we can rewrite the last two lines in the above expression as

$$\begin{aligned}
\mathbf{P}_t &= E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})'\right] - \mathbf{K}_t \mathrm{Cov}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] - \\
&\quad \mathbf{K}_t \mathrm{Var}\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\mathbf{K}_t' + \mathbf{K}_t \mathrm{Var}\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right)\mathbf{K}_t' \\
&= E_{t-1}\left[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})'\right] - \mathbf{K}_t \mathrm{Cov}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right].
\end{aligned}$$

Finally from subsection 13.3.2 we know $E_{t-1}\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)'\right] \equiv \mathbf{P}_{t|t-1}$. Accordingly the above equation can be written as

$$\begin{aligned}
\mathbf{P}_t &= \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathrm{Cov}\left[\left(\mathbf{y}_t - \mathbf{y}_{t|t-1}\right), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] \\
&= \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathrm{Cov}\left[\left(\mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t - \mathbf{y}_{t|t-1}\right), (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] \\
&= \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{Z}_t \mathrm{Cov}\left[\boldsymbol{\alpha}_t, (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] - \mathbf{K}_t \mathrm{Cov}\left[\mathbf{d}_t, (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] + \\
&\quad \mathbf{K}_t \mathrm{Cov}\left[\boldsymbol{\varepsilon}_t, (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] - \mathbf{K}_t \mathrm{Cov}\left[\mathbf{a}_{t|t-1}, (\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})\right] \\
&= \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{Z}_t \mathrm{Cov}\left(\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_t\right) + \mathbf{K}_t\mathbf{Z}_t \mathrm{Cov}\left(\boldsymbol{\alpha}_t, \mathbf{a}_{t|t-1}\right) \\
&= \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{Z}_t \mathrm{Var}\left(\boldsymbol{\alpha}_t\right) \\
&= \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{Z}_t\mathbf{P}_{t|t-1}
\end{aligned}$$

as defined in (13.4.6).

13.4. Maximum likelihood estimation

13.4.1. Unconstrained maximum likelihood

At time $t = 1, \ldots, T$, the Kalman filter provides the optimal estimate for the state vector. The joint density function of the T vectors of observations is given by

$$\ln L(\mathbf{y}, \boldsymbol{\psi}) = \prod_{t=1}^{N} p(\mathbf{y}_t|\mathbf{Y}_{t-1}), \tag{13.9}$$

where $\boldsymbol{\psi}$ is a vector of *hyperparameters* and $p(\mathbf{y}_t|\mathbf{Y}_{t-1})$ is the distribution of $\mathbf{y}_t$, conditional on the information set, $\mathbf{Y}$ at time $t-1$. Given the information set $\mathbf{Y}_{t-1}$, the true state vector is normally distributed with mean $\mathbf{a}_t$ and covariance matrix $\mathbf{P}_t$. Hence, $\mathbf{y}_t$ is also normally distributed with mean $\mathbf{y}_{t|t-1} = \mathbf{Z}_t\mathbf{a}_{t|t-1} + \mathbf{d}_t$ and error covariance matrix

$\mathbf{F}_t$. Accordingly the the log likelihood function is given by

$$\ln L(\mathbf{y}, \boldsymbol{\psi}) = -\frac{nT}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \ln |\mathbf{F}_t| - \frac{1}{2} \sum_{t=1}^{T} \boldsymbol{\nu}_t' \mathbf{F}_t \boldsymbol{\nu}_t, \tag{13.10}$$

which has to be maximized with respect to the vector of hyperparameters $\boldsymbol{\psi}$ (defined in Section 13.1)

13.4.2. Constrained maximum likelihood

To constrain the parameter estimates, the log likelihood function, $\ln L(\mathbf{y}, \boldsymbol{\psi})$ can be maximized subject to restrictions on the parameter vector, $\boldsymbol{\psi}$. The constraints can be expressed in equality $G(\boldsymbol{\psi}) = 0$ and inequality, $H(\boldsymbol{\psi}) \geq 0$ terms with the G and H functions encompassing both nonlinear and linear forms. Boundary restrictions, e.g. $\boldsymbol{\psi}_l \leq \boldsymbol{\psi} \leq \boldsymbol{\psi}_u$, where the subscripts l and u refer to the lower and upper bounds respectively, can simply be treated as a linear inequalities. Moreover, inequality constraints can be re-expressed at equality, i.e. $H(\boldsymbol{\psi}) \geq 0$ becomes $H(\boldsymbol{\psi}) = \zeta^2$ where ζ is a conformable vector of slack parameters. The Kuhn-Tucker (1951) conditions require that a distinction is made between active and inactive inequality constraints. Active constraints, $H_{\otimes}(\boldsymbol{\psi})$ have non-zero Lagrangians, γ_j and zero slack parameters ζ_j while the reverse holds for inactive constraints, $H_{\ominus}(\theta)$. To impose both equality and inequality restrictions the following augmented log-likelihood function would be maximized

$$L_A = \ln L(\mathbf{y}, \boldsymbol{\psi}) + \sum_{j=1}^{J} \lambda_j g(\boldsymbol{\psi})_j + \sum_{k=1}^{K} \gamma_k h_{\otimes}(\boldsymbol{\psi})_k + \sum_{l=1}^{L} h_{\ominus}(\boldsymbol{\psi})_l - \zeta_l^2 \tag{13.11}$$

where $\ln L(\mathbf{y}, \boldsymbol{\psi})$ is given by (13.11).

14. MATLAB programs by Peter Ireland

Peter Ireland has made MATLAB programs (i.e. both scripts and functions) available so the econometric work in his paper "A Method for Taking Models to the Data," in the March 2004 issue of the Journal of Economic Dynamics and Control can be replicated. These programs can be found at: http://www2.bc.edu/~irelandp/#programs. A brief taxonomy of the programs according to the tasks performed include:

Estimation

- **est.m:** Maximizes (minimizes negative) log-likelihood function for Hansen's RBC model to obtain the parameter estimates where some have been transformed to comply with theoretical restrictions.
- **estse.m:** Computes standard errors for the transformed parameter estimates.
- **estseq.m:** Recursive estimation of **est.m** with variable end dates.
- **llfn.m:** *MATLAB function* called by **est.m** and **estseq.m**. Uses the Kalman filter to evaluate the log-likelihood function based on the transformed parameters.
- **llfnse.m:** *MATLAB function* called by **estse.m**. Uses the Kalman filter to evaluate the log-likelihood function based on the untransformed parameters.

Solution & Simulation

- **solv.m:** Solves Hansen's RBC model.
- **check.m:** Checks the solution found by the script **solv.m**.
- **ksmooth.m:** Provides smoothed estimates of the shocks in Hansen's model (i.e. technology plus measurement errors). Also calculates the correlation between the technology shock and the measurement errors.

Stability Tests

- **stabtest.m:** Performs parameter stability tests for the Hansen model (break point occurs at 1973:1).

Impulse Responses & Variance Decompositions

- **imp.m:** Computes impulse responses for Hansen's model using the solutions provided by **solv.m.**

- **vardec.m:** Uses the parameter estimates and standard errors provided by **estse.m** to calculate variance decompositions and standard errors for Hansen's model.

- **vardecfn.m:** *MATLAB function* which does the same as **vardec.m** (less the standard errors) for a given set of parameter estimates.

Forecasting

- **fork.m:** Generates *k-step-ahead* forecasts from four competing models: Hansen's model with vector *AR(1)* and scalar *AR(1)* errors plus a *UVAR(1)* and a *UVAR(2)* model.

- **forkfn.m:** *MATLAB function* which uses the Kalman filter to generate *k-step-ahead* forecasts for the non-predetermined variables in Hansen's model.

Extra MATLAB Programs

- The MATLAB programs briefly described above and fully listed below are relevant for the estimation and analysis of Ireland's full model (i.e. with vector *AR(1)* measurement errors). His model can also be estimated with scalar *AR(1)* errors (i.e. by assuming that the covariance matrix of his measurement equations, D and the covariance matrix of his measurement equations errors, V_2 are diagonal). His MATLAB scripts **estdse.m** and **estseqd.m** perform the same tasks as the above described programs **estse.m** and **estseq.m** except V_2 and D are diagonal. This also applies to the MATLAB functions **llfnd.m** and **llfndse.m** which correspond to **llfn.m** and **llfnse.m**. The program **stabtst2.m** performs the same stability tests as **stabtest.m** except the break point is 1980:1. Finally the functions **llfn2.m** & **llfnse2.m** and **llfn3.m** & **llfnse3.m** correspond to **llfn.m** & **llfnse.m** except the time trend is adjusted to accommodate the post 1973 and post 1981 sub-samples respectively.

References

[1] Andrews, Donald, W.K. and Ray C. Fair (1988) "Inference in Nonlinear Econometric Models with Structural Change" *Review of Economic Studies*, 55: 615-39.

[2] Azariadies, Costas (1993) *Intertemporal Macroeconomics*, Blackwell.

[3] Blanchard, Olivier J. and Kahn, Charles M. (1980) "The Solution of Linear Difference Models Under Rational Expectations" *Econometrica* 48: 1305-11.

[4] Davidson, R. and J. Mackinnon (1993) Estimation and Inference in Econometrics, Oxford University Press.

[5] Diebold, Francis X. and Roberto S. Mariano (1995) "Comparing Predictive Accuracy" *Journal of Business and Economic Statistics*, 13: 253-63.

[6] Greene, William H. (2003), *Econometric Analysis*, 5th ed., Prentice Hall.

[7] Hall, George J. (1996) "Overtime Effort, and the Propagation of Business Cycle Shocks", *Journal of Monetary Economics*, 38: 607-26.

[8] Harvey, Andrew C. (1991) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

[9] Harvey, D., S. Leybourne, and P. Newbold (1997) "Testing the Equality of Prediction Mean Squared Errors" *International Journal of Forecasting* 13: 281-91.

[10] Ireland, Peter N. (1999) "A Method for Taking Models to the Data.", Working Paper 9903. Federal Reserve Bank of Cleveland (1999); *mimeo* (2003) and *Journal of Economic Dynamics and Control,24: 1205-26 (*2004).

[11] Judd, Kenneth L. (1998) *Numerical Methods in Economics,* MIT Press.

[12] Kuhn, H.W. and A.W. Tucker (1951) "Nonlinear Programming" in J. Neyman (ed.), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, California, 481-92.

[13] Leslie, Derek (1993) *Advanced Macroeconomics*, McGraw Hill.

[14] Lütkepohl, Helmut (1996), *Handbook of Matricies*, John Wiley and Sons.

[15] Lütkepohl, Helmut (1991), *Introduction to Multiple Time Series Analysis*, Springer.

[16] Malley, J., Philippopoulos, A. and U. Woitek (2004) "The Roles of Technology and Political Shocks in Generating Growth and Fluctuations", *mimeo.*

[17] McGrattan Ellen, R. (1994) "The Macroeconomic Effects of Distortionary Taxation", *Journal of Monetary Economics*, 33: 573-601.

[18] McGrattan, Ellen R., Rogerson, Richard and Randall Wright (1994) "An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy", *International Economic Review*, 38: 267-90.

[19] Ruge-Murcia, Francisco J. (2003) Methods to Estimate Dynamic Stochastic General Equilibrium Models, *mimeo*, University of Montreal.

[20] Sargent, Thomas J. (1989) "Two Models of Measurements and the Investment Accelerator", *Journal of Political Economy*, 97: 251-87.

[21] Schoenberg, Ronald (1997) "Constrained Maximum Likelihood" *Computational Economics*, 10: 251-66.

[22] Shone, Ronald (1997) *Economics Dynamics*, Cambridge University Press.

[23] Stuart, Alan, Ord, Keith and Steven Arnold (1999) *Kendall's Advanced Theory of Statistics: Classical Inference & the Linear Model*, (6th edition), Arnold Publishers.

[24] Watson, Mark W. (1993) "Measures of Fit for Calibrated Models", *Journal of Political Economy*, 101: 1011-41.

[25] Woitek, Ulrich (2004) Matlab Workshop, University of Zurich, *mimeo.*

[26] Woodford, M (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.