Introduction

A convenient way of representing an economic time series y_t is through the so-called *trend-cycle decomposition*

$$\begin{aligned} y_t &= TD_t + Z_t \\ TD_t &= \text{deterministic trend} \\ Z_t &= \text{ random cycle/noise} \end{aligned}$$

For simplicity, assume

$$\begin{aligned} TD_t &= \kappa + \delta t \\ \phi(L)Z_t &= \theta(L)\varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma^2) \end{aligned}$$

where $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$and $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$. It is assumed that the polynomial $\phi(z) = 0$ has at most one root on the complex unit circle and $\theta(z) = 0$ has all roots outside the unit circle.

Trend Stationary and Difference Processes

Defn: The series y_t is called *trend stationary* if the roots of $\phi(z) = 0$ are outside the unit circle

Defn: The series y_t is called *difference stationary* if $\phi(z) = 0$ has one root on the unit circle and the others outside the unit circle.

If y_t is trend stationary then $\phi(L)$ is invertible and Z_t has the stationary or Wold representation

$$\begin{aligned} Z_t &= \phi(L)^{-1}\theta(L)\varepsilon_t \\ &= \psi(L)\varepsilon_t \end{aligned}$$

where

$$\begin{aligned} \psi(L) &= \phi(L)^{-1}\theta(L) = \sum_{k=0}^{\infty} \psi_k L^k \\ \psi_0 &= 1 \text{ and } \psi(1) \neq 0. \end{aligned}$$

If y_t is difference stationary then $\phi(L)$ can be factored as

$$\phi(L) = (1 - L)\phi^*(L)$$

where $\phi^*(z) = 0$ has all $p-1$ roots outside the unit circle. In this case, ΔZ_t has the stationary ARMA$(p-1, q)$ representation

$$\begin{aligned} \Delta Z_t &= \phi^*(L)^{-1}\theta(L)\varepsilon_t \\ &= \psi^*(L)\varepsilon_t \end{aligned}$$

where

$$\begin{aligned} \psi^*(L) &= \phi^*(L)^{-1}\theta(L) = \sum_{k=0}^{\infty} \psi_k^* L^k \\ \psi_0^* &= 1 \text{ and } \psi^*(1) \neq 0. \end{aligned}$$

Example: Difference stationary AR(2)

Let

$$\phi(L)Z_t = \varepsilon_t,\ \phi(L) = 1 - \phi_1 L - \phi_2 L^2$$

Assume that $\phi(z) = 0$ has one root equal to unity and the other root real valued with absolute value less than 1. Factor $\phi(L)$ so that

$$\begin{aligned} \phi(L) &= (1 - \phi^* L)(1 - L) = \phi^*(L)(1 - L) \\ \phi^*(L) &= 1 - \phi^* L \text{ with } |\phi^*| < 1. \end{aligned}$$

Then

$$\phi(L)Z_t = (1 - \phi^* L)(1 - L)Z_t = (1 - \phi^* L)\Delta Z_t$$

so that ΔZ_t follows an AR(1) process.

I(1) and I(0) Processes

If the noise series Z_t is difference stationary then we say that Z_t is *integrated of order 1* and we write $Z_t \sim I(1)$. To see why, note that

$$\begin{aligned} \Delta Z_t &= \psi^*(L)\varepsilon_t = u_t \\ & u_t \text{ is stationary} \end{aligned}$$

It follows that $Z_t = Z_{t-1} + u_t$ and by recursive substitution starting at time $t = 0$ we have

$$Z_t = Z_0 + \sum_{k=1}^{t} u_k$$

so that Z_t can be represented as the (integrated) sum of t stationary innovations $\{u_k\}_{k=1}^{t}$.

Moreover, since u_t is stationary we say that u_t is *integrated of order zero*, and write $u_t \sim I(0)$, to signify that u_t cannot be written as the sum of stationary innovations.

Note: If ΔZ_t is an ARMA(p,q) process then Z_t is called an ARIMA(p,1,q) process. The term ARIMA refers to an autoregressive *integrated* moving average process.

Impulse Response Functions from I(1) Processes

Consider an $I(1)$ process with Wold representation $\Delta y_t = \psi^*(L)\varepsilon_t$. Since $\Delta y_t = y_t - y_{t-1}$ the level y_t may be represented as

$$y_t = y_{t-1} + \Delta y_t$$

Similarly, the level at time $t+h$ may be represented as

$$y_{t+h} = y_{t-1} + \Delta y_t + \Delta y_{t+1} + \cdots + \Delta y_{t+h}$$

The impulse response on the level of y_{t+h} of a shock to ε_t is

$$\begin{aligned} \frac{\partial y_{t+h}}{\partial \varepsilon_t} &= \frac{\partial \Delta y_t}{\partial \varepsilon_t} + \frac{\partial \Delta y_{t+1}}{\partial \varepsilon_t} + \cdots + \frac{\partial \Delta y_{t+h}}{\partial \varepsilon_t} \\ &= 1 + \psi_1^* + \cdots + \psi_h^* \end{aligned}$$

The long-run impact of a shock to the level of y_t is given by

$$\lim_{h\to\infty} \frac{\partial y_{t+h}}{\partial \varepsilon_t} = \sum_{j=1}^{\infty} \psi_j^* = \psi^*(1).$$

Hence, $\psi^*(1)$ measures the permanent effect of a shock, ε_t, to the level of y_t.

Remarks:

1. Since $\frac{\partial y_t}{\partial \varepsilon_t} = 1$ it follows that $\psi^*(1)$ can also be interpreted as the long-run effect of a shock relative to the immediate effect of a shock.

2. $\psi^*(1)$ is a natural measure of the importance of a permanent shock. If $\psi^*(1) = 1$ then the long-run effect of a shock is equal to the immediate effect; if $\psi^*(1) > 1$ the long-run effect is greater than the immediate effect; if $\psi^*(1) < 1$ the long-run effect is less than the immediate effect.

3. If $\psi^*(1) = 0$ then $Z_t \sim I(0)$. To see this suppose $Z_t \sim I(0)$ and has the Wold representation $Z_t = \psi(L)\varepsilon_t$ with $\psi(1) \neq 0$. Then

$$\begin{aligned} \Delta Z_t &= (1-L)Z_t = (1-L)\psi(L)\varepsilon_t = \psi^*(L)\varepsilon_t \\ \psi^*(L) &= (1-L)\psi(L) \end{aligned}$$

It follows that $\psi^*(1) = (1-1)\psi(1) = 0$.

Forecasting from an I(1) Process

Forecasting from an I(1) process follows directly from writing y_{t+h} as

$$y_{t+h} = y_t + \Delta y_{t+1} + \Delta y_{t+2} + \cdots + \Delta y_{t+h}$$

Then

$$\begin{aligned} y_{t+h|t} &= y_t + \Delta y_{t+1|t} + \Delta y_{t+2|t} + \cdots + \Delta y_{t+h|t} \\ &= y_t + \sum_{s=1}^{h} \Delta y_{t+s|t} \end{aligned}$$

Notice that forecasting an I(1) process proceeds from the most recent observation.

Example: Forecasting from an AR(1) model for Δy_t

Let Δy_t follow an AR(1) process

$$\Delta y_t - \mu = \phi(\Delta y_{t-1} - \mu) + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma^2)$$

where $|\phi| < 1$. Using the chain-rule of forecasting, the h-step ahead forecast of Δy_{t+h}based on information at time t is

$$\Delta y_{t+h|t} = \mu + \phi^h(\Delta y_t - \mu)$$

Then, the h-step ahead forecast of y_{t+h} is

$$\begin{aligned} y_{t+h|t} &= y_t + \sum_{s=1}^{h} [\mu + \phi^s(\Delta y_t - \mu)] \\ &= y_t + h\mu + (\Delta y_t - \mu) \sum_{s=1}^{h} \phi^h \end{aligned}$$

The Trend-Cycle Decomposition with Stochastic Trends

Assume $Z_t \sim I(1)$. Then it is possible to decompose Z_t into a stochastic (random walk) trend and a stationary, $I(0)$, "cyclical" component:

$$\begin{aligned} Z_t &= TS_t + C_t \\ & \quad TS_t \sim I(1) \\ & \quad C_t \sim I(0) \end{aligned}$$

The stochastic trend, TS_t, captures shocks that have a permanent effect on the level of y_t

The stationary component, C_t, captures shocks that only have a temporary effect on the level of y_t.

The components representation for y_t becomes

$$\begin{aligned} y_t &= TD_t + TS_t + C_t \\ TD_t + TS_t &= \text{ overall trend} \\ C_t &= \text{deviations about trend} \end{aligned}$$

Remarks:

1. The decomposition of Z_t into TS_t and C_t is not unique. In fact, there are an infinite number of such combinations depending on how TS_t and C_t are defined.

2. Two decompositions have been popular in the empirical literature: the Beveridge-Nelson (BN) decomposition; and the orthogonal unobserved components (UC0) decomposition. Both decompositions define TS_t as a pure random walk. They primarily differ in how they model the serial correlation in ΔZ_t.

3. The BN decomposition uses an unrestricted ARMA(p, q) model for ΔZ_t; the UC0 model uses a restricted ARMA(p, q) for ΔZ_t.

4. Recent work (e.g. Morley, Nelson, Zivot (2003)) has considered UC models with correlated components. Identification is tricky in these models

The Beveridge-Nelson Decomposition

Beveridge and Nelson (1980) proposed a definition of the permanent component of an $I(1)$ time series y_t with drift μ as the limiting forecast as horizon goes to infinity, adjusted for the mean rate of growth over the forecast horizon,

$$\begin{aligned} TD_t + BN_t &= \lim_{h\to\infty} \left(y_{t+h|t} - TD_{t+h|t} \right) \\ &= \lim_{h\to\infty} \left(y_{t+h|t} - \mu \cdot h \right) \end{aligned}$$

BN_t, is referred to as the *BN trend*.

The implied cycle at time t is then

$$C_t^{BN} = y_t - TD_t - BN_t$$

Beveridge and Nelson showed that if Δy_t has a Wold representation

$$\Delta y_t = \delta + \psi^*(L)\varepsilon_t$$

then BN_t follows a pure random walk without drift

$$\begin{aligned} BN_t &= BN_{t-1} + \psi^*(1)\varepsilon_t \\ &= BN_0 + \psi^*(1)\sum_{j=1}^{t} \varepsilon_t \end{aligned}$$

The derivation of the BN trend relies on the following algebraic result.

Let $\psi(L) = \sum_{k=0}^{\infty} \psi_k L^k$ with $\psi_0 = 1$. Then

$$\begin{aligned}
\psi(L) &= \psi(1) + (1 - L)\widetilde{\psi}(L), \\
\psi(1) &= \sum_{k=0}^{\infty} \psi_k, \\
\widetilde{\psi}(L) &= \sum_{j=0}^{\infty} \widetilde{\psi}_j L^j, \; \widetilde{\psi}_j = -\sum_{k=j+1}^{\infty} \psi_k.
\end{aligned}$$

In addition, if $\sum_{k=0}^{\infty} k|\psi_k| < \infty$ (1-summability) then $\sum_{k=0}^{\infty} |\widetilde{\psi}_k| < \infty$. 1-summability is satisfied by all covariance stationary ARMA(p,q) processes.

For an algebraic proof, see Hamilton (1993) pages 534 and 535.

Derivation of BN decomposition

Consider the Wold representation for Δy_t. By recursive substitution

$$y_t = y_0 + \delta t + \psi^*(L) \sum_{j=1}^{t} \varepsilon_j$$

Applying the decomposition $\psi^*(L) = \psi^*(1) + (1 - L)\widetilde{\psi}(L)$ gives

$$\begin{aligned} y_t &= y_0 + \delta t + \left(\psi^*(1) + (1-L)\widetilde{\psi}^*(L)\right) \sum_{j=1}^{t} \varepsilon_j \\ &= y_0 + \delta t + \psi^*(1) \sum_{j=1}^{t} \varepsilon_j + \widetilde{\varepsilon}_t - \widetilde{\varepsilon}_0 \\ &= TD_t + TS_t + C_t \end{aligned}$$

where

$$\begin{aligned} TD_t &= y_0 + \delta t \\ TS_t &= Z_0 + \psi^*(1) \sum_{j=1}^{t} \varepsilon_j \\ C_t &= \widetilde{\varepsilon}_t - \widetilde{\varepsilon}_0 \\ \widetilde{\varepsilon}_t &= \widetilde{\psi}^*(L)\varepsilon_t \end{aligned}$$

To show that $\psi^*(1)\sum_{j=1}^{t}\varepsilon_j$ is the BN trend, consider the series at time $t+h$

$$y_{t+h} = y_0 + \delta(t+h) + \psi^*(1)\sum_{j=1}^{t+h}\varepsilon_j + \widetilde{\varepsilon}_{t+h}$$

The forecast of y_{t+h} at time t is

$$y_{t+h|t} = y_0 + \delta(t+h) + \psi^*(1)\sum_{j=1}^{t}\varepsilon_j + \widetilde{\varepsilon}_{t+h|t}$$

The limiting forecast as horizon goes to infinity, adjusted for mean growth, is

$$\begin{aligned} \lim_{h\to\infty}\left(y_{t+h|t} - \delta h\right) &= y_0 + \delta t + \psi^*(1)\sum_{j=1}^{t}\varepsilon_j + \lim_{h\to\infty}\widetilde{\varepsilon}_{t+h|t} \\ &= y_0 + \delta t + \psi^*(1)\sum_{j=1}^{t}\varepsilon_j \\ &= TD_t + BN_t \end{aligned}$$

as $\lim_{h\to\infty}\widetilde{\varepsilon}_{t+h|t} = 0$ since $\widetilde{\varepsilon}_{t+h}$ is a mean-zero stationary process.

Example: BN decomposition from MA(1) process for Δy_t (Stock and Watson, 1987)

Let $y_t = \ln(\text{rgdp}_t)$. Using postwar quarterly data from 1947:II-1985:IV, Stock and Watson fit the following MA(1) model to the growth rate of real gdp:

$$\Delta y_t = 0.008 + \varepsilon_t + 0.3\varepsilon_{t-1},\ \varepsilon_t \sim iid(0, \sigma^2),\ \widehat{\sigma} = 0.0106$$

For the MA(1) model, the Wold representation for Δy_t has the simple form

$$\begin{aligned} \Delta y_t &= \delta + \psi^*(L)\varepsilon_t \\ \psi^*(L) &= 1 + \psi_1^* L,\ \psi_1^* = 0.03 \end{aligned}$$

Straightforward calculations give

$$\begin{aligned} \psi^*(1) &= 1 + \psi_1^* = 1.03 \\ \widetilde{\psi}_0^* &= -\sum_{j=1}^{\infty} \psi_j^* = -\psi_1^* = -0.03 \\ \widetilde{\psi}_j^* &= -\sum_{j=k+1}^{\infty} \psi_j^* = 0,\ j = 1, 2 \ldots \end{aligned}$$

The trend-cycle decomposition of y_t using the BN decomposition becomes

$$\begin{aligned} y_t &= (y_0 + \delta t) + \psi^*(1) \sum_{j=1}^{t} \varepsilon_j + \widetilde{\varepsilon}_t \\ &= y_0 + 0.008t + 1.3 \sum_{j=1}^{t} \varepsilon_j - 0.3\varepsilon_t \end{aligned}$$

so that

$$\begin{aligned} TD_t &= y_0 + 0.008t \\ BN_t &= 1.3 \sum_{j=1}^{t} \varepsilon_j \\ C_t &= -0.3\varepsilon_t \end{aligned}$$

Note that $\frac{\partial y_t}{\partial \varepsilon_t} = 1$ and $\frac{\partial y_{t+s}}{\partial \varepsilon_t} = 1.3$ for $s > 0$.

Remark: In the BN decomposition, $1.3\varepsilon_t$ is the shock to the trend and $-0.3\varepsilon_t$ is the shock to the cycle. It is tempting to conclude that the BN decomposition assumes that the trend and cycle shocks are perfectly negatively correlated. This is incorrect because ε_t only has the interpretation as a forecasting error.

The naive computation of the BN decomposition requires the following steps:

1. Estimation of ARMA(p,q) model for Δy_t

2. Estimation of $\psi^*(1)$ from estimated ARMA(p,q) model for Δy_t

3. Estimation of $\sum_{j=1}^{t} \varepsilon_j$ using residuals from estimated ARMA(p,q) model for Δy_t

Example: BN decomposition from ARMA(2,2) model for Δy_t

Morley, Nelson and Zivot (2003) fit the following ARMA(2,2) model to the growth rate of postwar quarterly real GDP over the period 1947:I - 1998:II

$$\begin{aligned}\Delta y_t &= 0.816 + 1.342\Delta y_{t-1} - 0.706\Delta y_{t-2} \\ &\quad + \hat{\varepsilon}_t - 1.054\hat{\varepsilon}_{t-1} + 0.519\hat{\varepsilon}_{t-2} \\ \phi(L) &= 1 - 1.342L + 0.706L^2, \\ \theta(L) &= 1 - 1.054L + 0.519L^2\end{aligned}$$

To compute an estimate of $\psi^*(1)$ from the ARMA(p,q) model, solve for the wold representation

$$\begin{aligned}\phi(L)\Delta y_t &= \theta(L)\varepsilon_t \\ &\Rightarrow \Delta y_t = \phi(L)^{-1}\theta(L)\varepsilon_t = \psi^*(L)\varepsilon_t\end{aligned}$$

where $\psi^*(L) = \phi(L)^{-1}\theta(L)$. Therefore,

$$\psi^*(1) = \phi(1)^{-1}\theta(1)$$

The estimate of $\psi^*(1)$ from the ARMA(2,2) model is

$$\psi^*(1) = \frac{1 - 1.054 + 0.519}{1 - 1.342 + 0.706} = 1.276$$

The estimate of the permanent component is then

$$TD_t + BN_t = y_0 + 0.816t + 1.276 \sum_{j=1}^{t} \hat{\varepsilon}_j$$

Example: BN decomposition from AR(1) process for Δy_t

From the formula for the $h-$step ahead forecast for y_t, it is easy to analytically compute the BN trend for y_t:

$$\begin{aligned} TD_t + BN_t &= \lim_{h\to\infty}\left(y_{t+h|t} - h\delta\right) \\ &= y_t + (\Delta y_t - \delta)\lim_{h\to\infty}\sum_{s=1}^{h}\phi^h \\ &= y_t + \frac{\phi}{1-\phi}(\Delta y_t - \delta) \end{aligned}$$

since

$$\begin{aligned} \lim_{h\to\infty}\sum_{s=1}^{h}\phi^h &= \sum_{s=0}^{\infty}\phi^h - 1 = \frac{1}{1-\phi} - 1 \\ &= \frac{\phi}{1-\phi} \end{aligned}$$

The cycle component is then

$$\begin{aligned} C_t^{BN} &= y_t - TD_t - BN_t \\ &= \frac{\phi}{1-\phi}(\Delta y_t - \mu) \end{aligned}$$

Morley (2002) shows how the BN decomposition for an AR(1) model for Δy_t may be extended to any model for Δy_t that can be represented in state space form. In particular, suppose $\Delta y_t - \mu$ is a linear combination of the elements of the $m \times 1$ state vector $\boldsymbol{\alpha}_t$:

$$\Delta y_t - \mu = \begin{bmatrix} z_1 & z_2 & \cdots & z_m \end{bmatrix} \boldsymbol{\alpha}_t$$

where z_i $(i = 1, \ldots, m)$ is the weight of the ith element of $\boldsymbol{\alpha}_t$ in determining $\Delta y_t - \mu$. Suppose further that

$$\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t, \ \boldsymbol{\eta}_t \sim \text{iid } N(\mathbf{0}, \mathbf{Q}),$$

such that all of the eigenvalues of $\mathbf{T}$ have modulus less than unity, and $\mathbf{T}$ is invertible. Then, Morley shows that

$$\begin{aligned} TD_t + BN_t &= \mathbf{y}_t + \begin{bmatrix} z_1 & z_2 & \cdots & z_m \end{bmatrix} \mathbf{T}(\mathbf{I}_m - \mathbf{T})^{-1}\mathbf{a}_{t|t} \\ C_t^{BN} &= y_t - TD_t - BN_t \\ &= -\begin{bmatrix} z_1 & z_2 & \cdots & z_m \end{bmatrix} \mathbf{T}(\mathbf{I}_m - \mathbf{T})^{-1}\mathbf{a}_{t|t} \end{aligned}$$

where $\mathbf{a}_{t|t}$ denotes the filtered estimate of $\boldsymbol{\alpha}_t$ from the Kalman filter recursions.

The Orthogonal Unobserved Components (UC) Model

The basic idea behind the UC model is to give structural equations for the components on the trend-cycle decomposition. For example, Watson (1986) considered UC-ARIMA models of the form

$$
\begin{aligned}
y_t &= \mu_t + C_t \\
\mu_t &= \alpha + \mu_{t-1} + \varepsilon_t, \ \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \\
\phi(L)C_t &= \theta(L)\eta_t, \ \eta_t \sim iid(0, \sigma_\eta^2) \\
\phi(L) &= 1 - \phi_1 L - \cdots - \phi_p L^p \\
\theta(L) &= 1 + \theta_1 L + \cdots + \theta_q L^q
\end{aligned}
$$

Identification:

1. The parameters of the UC model are not identified without further restrictions.

2. Restrictions commonly used in practice to identify all of the parameters are: (1) the roots of $\phi(z) = 0$ are outside the unit circle; (2) $\theta(L) = 1$, and (3) $cov(\varepsilon_t, \eta_t) = 0$. These restrictions identify C_t as a transitory autoregressive "cyclical" component, and μ_t as the permanent trend component.

3. The restriction $cov(\varepsilon_t, \eta_t) = 0$ states that shocks to C_t and μ_t are uncorrelated. As shown in Morley, Nelson and Zivot (2003), for certain models the assumption that $cov(\varepsilon_t, \eta_t) = 0$ turns out to be an over-identifying restriction.

Example: Clark's (1986) JPE Model

Clark considered the UC-ARIMA(2,0) model

$$\begin{aligned}
y_t &= \mu_t + C_t \\
\mu_t &= \alpha + \mu_{t-1} + \varepsilon_t,\ \varepsilon_t \sim \text{iid } N(0, \sigma_\varepsilon^2) \\
C_t &= \phi_1 C_{t-1} + \phi_2 C_{t-2} + \eta_t, \eta_t \sim \text{iid } N(0, \sigma_\eta^2) \\
cov(\varepsilon_t, \eta_t) &= 0
\end{aligned}$$

State Space Representation 1

Define $\boldsymbol{\alpha}_t = (\mu_t, C_t, C_{t-1})'$. Then the transition equation is

$$\begin{aligned}
\begin{pmatrix} \mu_t \\ C_t \\ C_{t-1} \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ C_{t-1} \\ C_{t-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \\
\mathbf{Q}_t &= \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix}
\end{aligned}$$

The measurement is

$$y_t = (1, 1, 0) \left(\begin{array}{c} \mu_t \\ C_t \\ C_{t-1} \end{array} \right)$$

Notice that

$$\begin{array}{rcl} \alpha_{1,t} & = & \mu_t \sim I(1) \\ \boldsymbol{\alpha}'_{2,t} & = & (C_t, C_{t-1})' \sim I(0) \end{array}$$

As a result, the distribution of the initial state vector $\boldsymbol{\alpha}_0 = (\mu_0, C_0, C_{-1})'$ cannot be determined in the usual way. This is because, $a_{1,0}$ does not have a simple stationary unconditional distribution. Since $var(\mu_t) = \mu_0 + \sigma_\varepsilon^2 \cdot t$, the usual approach is to assume

$$\begin{array}{rcl} a_{1,0} & = & 0 \\ var(\alpha_{1,0}) & = & \kappa \cdot 10^6 \end{array}$$

The distribution for $\boldsymbol{\alpha}_{2,0}$ may be based on its unconditional stationary distribution. Therefore, the initial state distribution is characterized by

$$\begin{array}{rcl} \mathbf{a}_0 & = & \mathbf{0} \\ \mathbf{P}_0 & = & \left(\begin{array}{cc} \kappa \cdot 10^6 & 0 \\ 0 & \mathbf{P}_2 \end{array} \right) \end{array}$$

State Space Representation 2

Since $y_t \sim I(1)$, consider the transformed model

$$\begin{aligned} \Delta y_t &= \Delta\mu_t + \Delta C_t \\ &= \alpha + \Delta C_t + \varepsilon_t \end{aligned}$$

Define $\boldsymbol{\alpha}_t = (C_t, C_{t-1})'$. Then the transition equation is

$$\begin{pmatrix} C_t \\ C_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C_{t-1} \\ C_{t-2} \end{pmatrix} + \begin{pmatrix} \eta_t \\ 0 \end{pmatrix}$$

The corresponding measurement equation is

$$\Delta y_t = \begin{pmatrix} 1 & -1 \end{pmatrix} \boldsymbol{\alpha}_t + \varepsilon_t$$

Since $\boldsymbol{\alpha}_t$ is covariance stationary, the distribution of the initial state vector $\boldsymbol{\alpha}_0$ may be determined in the usual way.

Proposition: Any UC-ARIMA model with a specified correlation between ε_t and η_t is observationally equivalent to an ARMA model for Δy_t with nonlinear restrictions on the parameters. The restricted ARMA model for Δy_t is called the *reduced form* of the UC-ARIMA model.

Proposition: (Lippi and Reichlin (1992) JME) $\psi^*(1)$ computed from the reduced form ARMA model for Δy_t based on a UC-ARIMA model for y_t is always less than 1.

Example: Random walk plus noise model

Consider the random walk plus noise model

$$\begin{aligned} y_t &= \mu_t + \eta_t,\ \eta_t \sim \text{iid}(0, \sigma^2_\eta) \\ \mu_t &= \mu_{t-1} + \varepsilon_t\ \varepsilon_t \sim \text{iid}(0, \sigma^2_\varepsilon) \\ q &= \frac{\sigma^2_\varepsilon}{\sigma^2_\eta} = \text{signal-to-noise ratio} \\ & \eta_t \text{ and } \varepsilon_t \text{ are independent} \end{aligned}$$

The reduced form ARMA model for Δy_t is

$$\Delta y_t = \Delta \mu_t + \Delta \eta_t = \varepsilon_t + \eta_t - \eta_{t-1}.$$

Claim: Δy_t follows an ARMA(0,1) process. To see this, consider the autocovariances of Δy_t :

$$\begin{aligned} \gamma^*_0 &= var(\Delta y_t) = var(\varepsilon_t + \eta_t - \eta_{t-1}) = \sigma^2_\varepsilon + 2\sigma^2_\eta \\ &= \sigma^2_\eta (q+2) \\ \gamma^*_1 &= cov(\Delta y_t, \Delta y_{t-1}) \\ &= E[(\varepsilon_t + \eta_t - \eta_{t-1})(\varepsilon_{t-1} + \eta_{t-1} - \eta_{t-2})] \\ &= -\sigma^2_\eta \\ \gamma^*_j &= 0,\ j > 1 \end{aligned}$$

The autocovariances for Δy_t are the same as the autocovariances for an ARMA(0,1) model, and so Δy_t has the representation

$$\Delta y_t = \varsigma_t + \theta\varsigma_{t-1},\ \varsigma_t \sim \text{iid}(0, \sigma_\varsigma^2)$$

where the ARMA(0,1) parameters θ and σ_ς^2 are nonlinearly related to the UC-ARIMA parameters σ_η^2 and σ_ε^2, and the error term ς_t encompasses the structural shocks η_t and ε_t.

Identification of structural parameters

1. The reduced form ARMA(0,1) model has two parameters θ and σ_{ς}^2

2. The structural random walk plus noise model also has two parameters σ_{ε}^2 and σ_{η}^2.

3. Since the reduced form and structural models have the same number of parameters the order condition for identification is satisfied.

4. If ε_t and η_t were allowed to be correlated then there would be three structural parameters (two variances and a covariance) and only two reduced form parameters, and the order condition for identification would not be satisfied. Hence, setting $cov(\varepsilon_t, \eta_t) = 0$ is an identifying restriction in this model.

The mapping from UC-ARIMA parameters to the reduced form parameters is determined as follows. The first order autocorrelation for the reduced form ARMA(0,1) is

$$\rho_1^* = \frac{\theta}{1+\theta^2}$$

and for the UC-ARIMA it is

$$\rho_1 = \frac{\gamma_1^*}{\gamma_0^*} = \frac{-1}{q+2}$$

Setting $\rho_1^* = \rho_1$ and solving for θ gives

$$\theta = \frac{-(q+2) \pm \sqrt{(q+2)^2 - 4}}{2}$$

The invertible solution is

$$\theta = \frac{-(q+2) + \sqrt{q^2+4q}}{2}, \ \theta < 0.$$

Remarks:

1. Notice that the MA coefficient θ for the reduced form ARMA(0,1) model is restricted to be negative.

2. If $q = 0$, then $\theta = -1$. Similarly, matching variances for the two models gives

$$\sigma_\varsigma^2 = \frac{\sigma_\eta^2}{-\theta}.$$

3. Since $\theta < 0$, it follows that

$$\psi^*(1) = 1 + \theta < 1.$$