# Difference equations

**Definitions**:

A *difference equation* takes the general form

$$x_t = f(x_{t-1}, x_{t-2}, \dots)$$

defining the current value of a variable $x$ as a function of previously generated values.

A *finite order* ($m$th order) difference equation takes the general form

$$x_t = f(x_{t-1}, \dots, x_{t-m}).$$

A *linear* difference equation takes the general form

$$x_t = \alpha + \lambda_1 x_{t-1} + \lambda_2 x_{t-2} + \dots$$

A *stochastic* difference equation takes the general form

$$x_t = f(x_{t-1}, x_{t-2}, \dots, \varepsilon_t)$$

where $\{\varepsilon_t\}$ is a random sequence (often i.i.d. in applications) called the ‘forcing process’ or ‘driving process’.

A *linear stochastic* difference equation takes the general form

$$x_t = \alpha + \lambda_1 x_{t-1} + \lambda_2 x_{t-2} + \ldots + \varepsilon_t$$

The object is to solve these equations (determine path $x_1, x_2, x_3, \ldots$ ) given initial conditions $x_0, x_{-1}, \ldots$ and sequence $\{\varepsilon_t\}$.

Consider the non-stochastic case to establish methods and notation. The linear case is the simplest and best understood, and we focus on this.

# Some useful formulae

A geometric series is

$$A = 1 + a + a^2 + a^3 + \cdots = \sum_{i=0}^{\infty} a^i.$$

Let the $t$th partial sum be denoted

$$A(t) = 1 + a + a^2 + \cdots + a^{t-2} + a^{t-1} = \sum_{i=0}^{t-1} a^i.$$

Suppose $a \neq 1$. Then, note that

$$aA(t) = a + a^2 + \cdots + a^{t-1} + a^t = A(t) + a^t - 1$$

which solves as

$$A(t) = aA(t) + 1 - a^t = \frac{1 - a^t}{1 - a}.$$

Cases:

1. If $|a| < 1$, the geometric series is *convergent*. $a^t \to 0$ as $t \to \infty$, and
$$A = \lim_{t\to\infty} A(t) = \frac{1}{1-a}$$

2. If $a > 1$ it *diverges*: $A = \infty$.

3. If $a = -1$, no solution. $A(t)$ "flip-flops" between 0 and 1.

4. If $a < -1$, $A(t)$ ‘flip-flops’ between $\pm\infty$ in limit!

5. Finally, if $a = 1$, $A(t) = t$ and $A = \infty$.

Also consider

$$A^*(t) = a + 2a^2 + 3a^3 + \cdots + ta^t$$

By a similar argument,

$$\begin{aligned}(1-a)A^*(t) &= a + 2a^2 + 3a^3 + \cdots + ta^t - (a^2 + 2a^3 + \cdots + ta^{t+1}) \\ &= a + a^2 + a^3 + \cdots + a^t - ta^{t+1}\end{aligned}$$

Hence,

$$A^*(t) = \frac{a}{1-a}(A(t) - ta^t).$$

When $|a| < 1$, note that

$$A^* = \sum_{i=1}^{\infty} ia^i = \frac{a}{1-a}A = \frac{a}{(1-a)^2}$$

# First order linear difference equation

$$x_t = \alpha + \lambda_1 x_{t-1}$$

Given $x_0$, the solution path is found by iteration, as

$$x_1 = \alpha + \lambda_1 x_0$$

$$x_2 = \alpha(1 + \lambda_1) + \lambda_1^2 x_0$$

$$\cdots$$

$$x_t = \alpha(1 + \lambda_1 + \cdots + \lambda_1^{t-1}) + \lambda_1^t x_0$$

If $|\lambda_1| < 1$ the series is summable, and as $t \to \infty$ ,

$$x_t \to \alpha \sum_{j=0}^{\infty} \lambda_1^j = \frac{\alpha}{1 - \lambda_1}.$$

This is called the *stable solution*, and is independent of $x_0$. $x_t$ approaches this point from any starting point.

In the other cases of $\lambda_1$ , there is either no stable solution, or an infinite solution.

Note: we can guess the solution by putting $x_t = x_{t-1} = x$ (say) and so solve

$$x = \frac{\alpha}{1 - \lambda_1}$$

This must be the stable solution if one exists – but otherwise, it is irrelevant.

# Second Order Linear Difference Equation

$$x_t = \alpha + \lambda_1 x_{t-1} + \lambda_2 x_{t-2}$$

Our concern is to find conditions for a stable solution to this equation, as in the first-order case. *If it exists*, this must take the form

$$x = \frac{\alpha}{1 - \lambda_1 - \lambda_2}$$

However, solution by iteration is obviously difficult:

$$\begin{aligned} x_t &= \alpha + \lambda_1 x_{t-1} + \lambda_2 x_{t-2} \\ &= \alpha(1 + \lambda_1 + \lambda_2) + (\lambda_1^2 + \lambda_2)x_{t-2} + \lambda_1\lambda_2 x_{t-3} \\ &= \cdots ? \end{aligned}$$

How do we know if the solution converges?

Alternatively, consider the pair of first-order equations

$$x_t = \mu_1 x_{t-1} + y_t$$
$$y_t = \alpha + \mu_2 y_{t-1}.$$

Write these in the form

$$\begin{aligned}\alpha &= (x_t - \mu_1 x_{t-1}) - \mu_2(x_{t-1} - \mu_1 x_{t-2}) \\ &= x_t - (\mu_1 + \mu_2)x_{t-1} + \mu_1\mu_2 x_{t-2} \\ &= x_t - \lambda_1 x_{t-1} - \lambda_2 x_{t-2}\end{aligned}$$

It is intuitively clear that the stability conditions have the form $|\mu_1| < 1,\ |\mu_2| < 1$, since then

$$y_t \to \frac{\alpha}{1-\mu_2}$$

and

$$x_t \to \frac{1}{1-\mu_1}\left(\frac{\alpha}{1-\mu_2}\right) = \frac{\alpha}{1-\lambda_1-\lambda_2}$$

.

# The Lag Operator

To apply these restrictions, it is necessary to invert the mapping $(\mu_1, \mu_2) \to (\lambda_1, \lambda_2)$.

Let the operator $L$ (alternative notation, $B$) be defined by

$$Lx_t = x_{t-1}.$$

Then, for example,

$$L^2 x_t = L(Lx_t) = Lx_{t-1} = x_{t-2}.$$

The second-order equation can now be written

$$(1 - \mu_1 L)x_t = y_t$$

$$(1 - \mu_2 L)y_t = \alpha$$

or equivalently,

$$\begin{aligned} \alpha &= (1 - \mu_1 L)(1 - \mu_2 L)x_t \\ &= (1 - (\mu_1 + \mu_2)L + \mu_1\mu_2 L^2)x_t \\ &= (1 - \lambda_1 L - \lambda_2 L^2)x_t. \end{aligned}$$

The model therefore involves a *quadratic equation in the lag operator*.

# Reminder: Roots of a Quadratic

Consider, for $z \in C$, the quadratic equation

$$z^2 - \lambda_1 z - \lambda_2 = (z - \mu_1)(z - \mu_2) = 0$$

where $\lambda_1 = \mu_1 + \mu_2$ and $\lambda_2 = -\mu_1\mu_2$.

$\mu_1$ and $\mu_2$ are the roots (zeros) of this equation, and are given by

$$\mu_1 = \frac{-\lambda_1 + \sqrt{\lambda_1^2 - 4\lambda_2}}{2}, \quad \mu_2 = \frac{-\lambda_1 - \sqrt{\lambda_1^2 - 4\lambda_2}}{2}$$

- When $\lambda_2 > \lambda_1^2/4$ these solutions are complex numbers (complex conjugate pair, since $\lambda_1$ and $\lambda_2$ real).
- $y_t$ can be complex, although $x_t$ is real by construction.

# Stability Analysis

The stable solution (finite, independent of initial conditions) evidently takes the form

$$x_t = \frac{\alpha}{1 - \lambda_1 L - \lambda_2 L^2} \to \frac{\alpha}{1 - \lambda_1 - \lambda_2}$$

*provided* the inversion of the lag polynomial is a legitimate step.

Assume $|\mu_1| < 1$, $|\mu_2| < 1$ and consider for $z \in C$,

$$\frac{1}{1 - \lambda_1 z - \lambda_2 z^2} = \frac{1}{(1 - \mu_1 z)(1 - \mu_2 z)}.$$

Note that, for $|z| \leq 1$,

$$(1 - \mu_1 z)(1 + \mu_1 z + \mu_1^2 z^2 + \mu_1^3 z^3 + \cdots) = 1$$

so that

$$\frac{1}{1 - \mu_1 z} = \sum_{j=0}^{\infty} \mu_1^j z^j.$$

1. Assume $\mu_1 \neq \mu_2$ (both real). Then, for $|z| \leq 1$,

$$\frac{1}{(1 - \mu_1 z)(1 - \mu_2 z)} = \frac{1}{\mu_1 - \mu_2}\left(\frac{\mu_1}{1 - \mu_1 z} - \frac{\mu_2}{1 - \mu_2 z}\right) = \sum_{j=0}^{\infty}\left(\frac{\mu_1^{j+1} - \mu_2^{j+1}}{\mu_1 - \mu_2}\right) z^j$$

2. Assume $\mu_1 = \mu_2$.

Write $\mu_2 = \mu_1 + \delta$ in $\dfrac{\mu_1^{j+1} - \mu_2^{j+1}}{\mu_1 - \mu_2}$

By L'Hôpital's rule*,

$$\frac{(\mu_1+\delta)_1^{j+1} - \mu_1^{j+1}}{\delta} \to (j+1)\mu_1^j \quad \text{as } \delta \to 0,$$

and hence

$$\frac{1}{(1-\mu_1 z)^2} = \sum_{j=0}^{\infty} (j+1)\mu_1^j z^j$$

(Compare page 2.5 above.)

---

* If $f(\delta) \to 0$ and $g(\delta) \to 0$ as $\delta \to 0$, the limit of $\dfrac{f(\delta)}{g(\delta)}$ is equal to that of $\dfrac{f'(\delta)}{g'(\delta)}$, when the latter is defined.

# Complex Roots

Let $\mu_1 = re^{i\theta}$ and $\mu_2 = re^{-i\theta}$ (complex conjugate pair) where $i = \sqrt{-1}$ and $0 \leq \theta \leq 2\pi$.

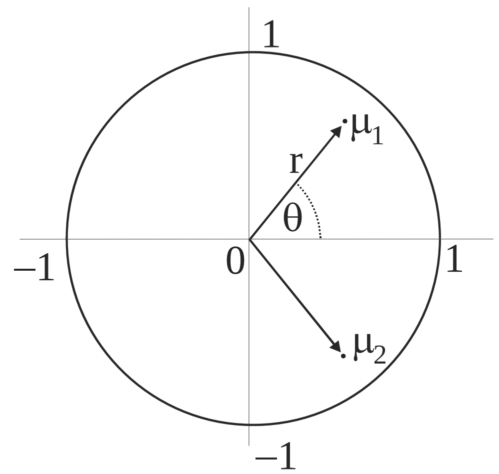
- Their modulus (absolute value) is

$$|\mu_1| = |\mu_2| = \sqrt{\mu_1\mu_2} = r.$$

- Using the facts $re^{i\theta} = r(\cos\theta + i\sin\theta)$, $\cos(-x) = \cos x$ and $\sin(-x) = -\sin(x)$, we obtain

$$\frac{\mu_1^{j+1} - \mu_2^{j+1}}{\mu_1 - \mu_2} = r^j \frac{e^{i(j+1)\theta} - e^{-i(j+1)\theta}}{e^{i\theta} - e^{-i\theta}} = r^j \frac{\sin(j+1)\theta}{\sin\theta}, \; j = 1, 2, 3, \ldots$$

- Stability depends on the roots lying inside the unit circle, having $r < 1$.



Finally..., replace $z$ by the operator $L$.

Since $L^j\alpha = \alpha$ for any $j$, we have the result

$$x_t \to \frac{\alpha}{1 - \lambda_1 - \lambda_2}$$

$$= \begin{cases} \alpha \sum_{j=0}^{\infty} \left( \frac{\mu_1^{j+1} - \mu_2^{j+1}}{\mu_1 - \mu_2} \right) & \text{real roots, } |\mu_1|,\ |\mu_2| < 1 \\ \alpha \sum_{j=0}^{\infty} (j+1)\mu_1^j & \text{equal roots, } |\mu_1| < 1 \\ \alpha \sum_{j=0}^{\infty} r^j \frac{\sin(j+1)\theta}{\sin\theta} r^j & \text{complex roots, } r < 1 \end{cases}$$

# Higher Order Cases

These are dealt with in the same way.

Since the lag coefficients are real, the roots must be either real, or in conjugate complex pairs.

Factorise the polynomial as

$$z^p - \lambda_1 z^{p-1} - \cdots - \lambda_p = (z-\mu_1)(z-\mu_2)\cdots(z-\mu_p).$$

The equation

$$x_t = \alpha + \lambda_1 x_{t-1} + \cdots + \lambda_p x_{t-p}$$

has a stable solution if $|\mu_k| < 1$ for $k = 1,\ldots,p$.

In particular, if $(\mu_k, \mu_{k+1})$ form a complex conjugate pair, then they lie must inside the unit circle, with $r < 1$.

# Alternative Representation

An equivalent approach is to let $w = 1/z$ and consider the roots of the equation

$$\begin{aligned}(w^{-2} - \lambda_1 w^{-1} - \lambda_2)w^2 &= 1 - \lambda_1 w - \lambda_2 w^2 \\ &= (1 - \mu_1 w)(1 - \mu_2 w) = 0.\end{aligned}$$

- The roots of this equation are $1/\mu_1$ and $1/\mu_2$.
- Thus, some authors give the stability condition as "the roots of the lag polynomial must lie *outside* the unit circle".
- Both usages are equally valid – important not to get confused!

# Stochastic Linear Difference Equations

First Order Case:

$$x_t = \alpha + \lambda_1 x_{t-1} + \varepsilon_t, \quad t = 1, 2, 3, \dots$$

The iterative solution is

$$x_t = \alpha(1 + \lambda_1 + \cdots + \lambda_1^{t-1}) + \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \cdots + \lambda_1^{t-1}\varepsilon_1 + \lambda_1^t x_0$$

Assume that $\varepsilon_t \sim iid(0, \sigma^2)$.

The 'solution' of this equation is interpreted in terms of the properties of the random variable $x_t$ when $t$ is large.

Case: $|\lambda_1| < 1$.

As $t \to \infty$ , dependence on starting value becomes negligible.

$$E(x_t) \to \frac{\alpha}{1 - \lambda_1} + \sum_{j=1}^{\infty} \lambda_1^j E(\varepsilon_{t-j}) = \frac{\alpha}{1 - \lambda_1}$$

$$\begin{aligned} \mathrm{Var}(x_t) \to E\left(\sum_{j=0}^{\infty} \lambda_1^j \varepsilon_{t-j}\right)^2 &= \sum_{j=0}^{\infty} \lambda_1^{2j} E(\varepsilon_{t-j}^2) + \text{terms in } E(\varepsilon_{t-j}\varepsilon_{t-k}) \text{ for } j \neq k \\ &= \sigma^2 \sum_{j=0}^{\infty} \lambda_1^{2j} = \frac{\sigma^2}{1 - \lambda_1^2}. \end{aligned}$$

# Other Properties

Since the forcing process is i.i.d., and $x_t$ depends (in effect) on only a finite number of these terms, it is a stationary process when $t$ is large enough.

$$\begin{aligned} x_{t+m} &= \frac{\alpha}{1-\lambda_1} + \sum_{j=0}^{\infty} \lambda_1^j \varepsilon_{t+m-j} \\ &= \frac{\alpha}{1-\lambda_1} + \lambda_1^m \left( \sum_{j=0}^{\infty} \lambda_1^j \varepsilon_{t-j} \right) + \sum_{j=0}^{m-1} \lambda_1^j \varepsilon_{t+m-j} \\ &= \frac{\alpha(1-\lambda_1^m)}{1-\lambda_1} + \lambda_1^m x_t + \sum_{j=0}^{m-1} \lambda_1^j \varepsilon_{t+m-j} \end{aligned}$$

and $E(x_t \varepsilon_{t+m-j}) = 0$ for $j < m$.

Hence,

$$\mathrm{Cov}(x_t, x_{t+m}) = \lambda_1^m \mathrm{Var}(x_t) = \frac{\lambda_1^m \sigma^2}{1-\lambda_1^2} = \gamma_m.$$

Therefore, $x_t$ is a short memory process, since

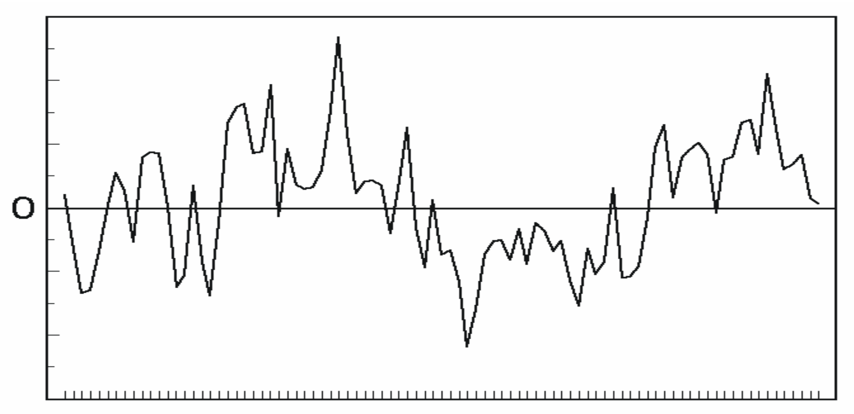
$$\sum_{j=0}^{\infty} |\gamma_j| = \frac{\sigma^2}{(1-\lambda_1^2)(1-|\lambda_1|)} < \infty.$$

Note that we may define a stationary process with starting point $t = 0$, by letting $x_0$ be a drawing from the stationary distribution of $x_t$.

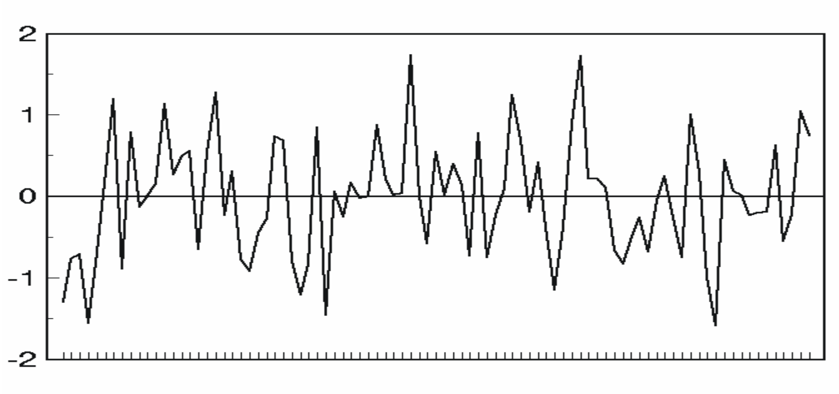
This stochastic process is called a first-order autoregression (AR(1)).

Realization of 100 observations from

$$x_t = 0.7x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad x_0 = 0$$




Corresponding i.i.d. process, $\varepsilon_t$:



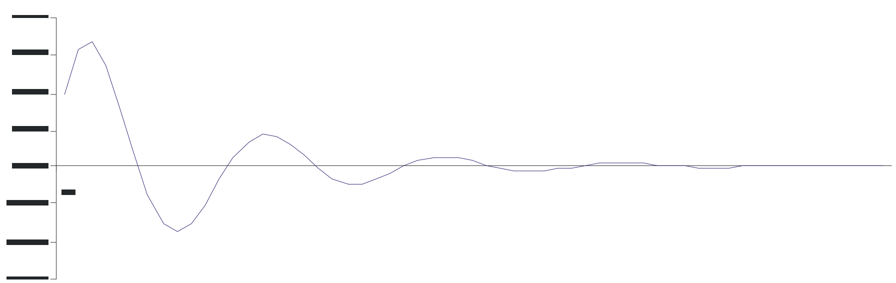
# Second Order Case.

$$x_t = \alpha + \lambda_1 x_{t-1} + \lambda_1 x_{t-2} + \varepsilon_t, \quad t = 1, 2, 3, \ldots$$

From previous results, we have the stationary second order solution (MA(∞) representation),

$$x_t \to \frac{\alpha + \varepsilon_t}{1 - \lambda_1 L - \lambda_2 L^2} = \begin{cases} \sum_{j=0}^{\infty} \left( \dfrac{\mu_1^{j+1} - \mu_2^{j+1}}{\mu_1 - \mu_2} \right) (\alpha + \varepsilon_{t-j}) & \text{real roots, } |\mu_1|, \ |\mu_2| < 1 \\ \sum_{j=0}^{\infty} (j+1)\mu_1^j (\alpha + \varepsilon_{t-j}) & \text{equal roots, } |\mu_1| < 1 \\ \sum_{j=0}^{\infty} r^j \dfrac{\sin(j+1)\theta}{\sin\theta} r^j (\alpha + \varepsilon_{t-j}) & \text{complex roots, } r < 1 \end{cases}$$

Complex roots imply sinusoidal lag distributions in the MA(∞) representation!



# Generalization

The AR($p$) process is

$$x_t = \alpha + \lambda_1 x_{t-1} + \cdots + \lambda_p x_{t-p} + \varepsilon_t.$$

Using the lag operator, this can be written in the form

$$x_t - \lambda_1 L x_t - \lambda_2 L^2 x_t - \cdots \lambda_p L^p x_t = \alpha + \varepsilon_t$$

or

$$\lambda(L) x_t = \alpha + \varepsilon_t$$

where

$$\lambda(L) = 1 - \lambda_1 L - \lambda_2 L^2 - \cdots - \lambda_p L^p.$$

The stationary solution of the model, when it exists, can be written in the form

$$x_t = \frac{\alpha}{\lambda(1)} + \frac{\varepsilon_t}{\lambda(L)}$$

where $1/\lambda(L)$ is a lag polynomial of infinite order, with summable coefficients.

# Autocovariances of AR processes

These are conveniently found from the *Yule-Walker equations*.

Multiply the equation by $x_{t-j}$ for $j = 0, 1, 2, \ldots, p$ and take expected values to yield a system of $p + 1$ equations in the unknowns $\gamma_0, \ldots \gamma_p$.

Consider the AR(2):

$$\gamma_0 - \lambda_1\gamma_1 - \lambda_2\gamma_2 = \sigma^2$$

$$\gamma_1 - \lambda_1\gamma_0 - \lambda_2\gamma_1 = 0$$

$$\gamma_2 - \lambda_1\gamma_1 - \lambda_2\gamma_0 = 0$$

These equations may be solved as

$$\begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 1 & -\lambda_1 & -\lambda_2 \\ -\lambda_1 & 1 - \lambda_2 & 0 \\ -\lambda_2 & -\lambda_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{\sigma^2}{(1 - \lambda_2)(1 - \lambda_2^2) - (1 + \lambda_2)\lambda_1^2} \begin{bmatrix} 1 - \lambda_2 \\ \lambda_1 \\ \lambda_2 + \lambda_1^2 - \lambda_2^2 \end{bmatrix}$$

For the higher order cases, solve the difference equation

$$\gamma_j = \lambda_1 \gamma_{j-1} + \lambda_2 \gamma_{j-2}$$

for $j = 3, 4, 5, \ldots$

Obviously, the conditions for $\gamma_j \to 0$ are identical to the stability conditions for the process.

- By rearranging the Y-W equations, one can also solve for $(\sigma^2, \lambda_1, \lambda_2)$ from $(\gamma_0, \gamma_1, \gamma_2)$.

# Moving Average Processes

Consider a process of the form

$$x_t = \alpha + \theta(L)\varepsilon_t$$

where $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$ and $\varepsilon_t \sim iid(0, \sigma^2)$.

This is the MA($q$) process.

The autocovariances are

$$\begin{aligned}
\gamma_0 &= \sigma^2(1 + \theta_1^2 + \cdots + \theta_q^2) \\
\gamma_1 &= \sigma^2(\theta_1 + \theta_2\theta_1 + \cdots + \theta_q\theta_{q-1}) \\
\gamma_2 &= \sigma^2(\theta_2 + \theta_3\theta_1 + \cdots + \theta_q\theta_{q-2}) \\
&\dots \\
\gamma_q &= \sigma^2\theta_q
\end{aligned}$$

with $\gamma_j = 0$ for $j > q$. Thus, the process is stationary for all choices of $\theta(L)$.

If $\theta(L)$ is invertible, it can be expressed as a difference equation of infinite order,

$$\frac{x_t}{\theta(L)} = \frac{\alpha}{\theta(1)} + \varepsilon_t.$$

All MA($q$) processes can be written as AR($\infty$) except those having a root of unity (over-differenced processes).

# Invertibility of MA Processes

Consider the MA(1) case first.

$$v_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

is a process with the following properties:

$$E(v_t) = 0, \quad \mathrm{Var}(v_t) = \sigma^2(1+\theta_1^2), \quad \mathrm{Cov}(v_t, v_{t-1}) = \theta_1 \sigma^2$$
$$\mathrm{Cov}(v_t, v_{t-j}) = 0 \text{ for } j > 1.$$

But suppose $\eta_t \sim iid(0, \omega^2)$ where $\omega^2 = \sigma^2\theta_1^2$. Then we can write *equivalently*

$$v_t' = \eta_t + \theta_1^{-1}\eta_{t-1}$$

- The processes $v_t$ and $v_t'$ have the same autocovariances, and on this basis are observationally equivalent.
- There is no loss of generality in always choosing the representation with $|\mu_1| < 1$.

The MA($q$) case:

$\theta(L)\varepsilon_t$ and $\theta^*(L)\eta_t$ are observationally equivalent processes where

$$\theta(L) = (1 - \mu_1 L)\cdots(1 - \mu_q L)$$

$$\theta^*(L) = (1 - \mu_1^{-1} L)(1 - \mu_2 L)\cdots(1 - \mu_q L)$$

and

$$E(\eta_t^2) = \mu_1^2 E(\varepsilon_t^2).$$

In total, $2^q$ equivalent representations!

Conventionally, we impose invertibility to *identify* the model.

# ARMA Processes

Combining AR and MA components yields the ARMA(p,q) process

$$\lambda(L)x_t = \alpha + \theta(L)\varepsilon_t.$$

This represents a flexible class of linear models for stationary processes.

Subject to stability/invertibility, the ARMA can be viewed as a difference equation of infinite order (AR($\infty$ )),

$$\frac{\lambda(L)}{\theta(L)}x_t = \frac{\alpha}{\theta(1)} + \varepsilon_t$$

and as a moving average of infinite order (MA($\infty$)):

$$x_t = \frac{\alpha}{\lambda(1)} + \frac{\theta(L)}{\lambda(L)}\varepsilon_t.$$

# Vector Autoregressions

Let $\{\boldsymbol{x}_t\}$ be a sequence of $m$-vectors. The VAR($p$) model takes the form

$$\boldsymbol{x}_t = \boldsymbol{a}_0 + \boldsymbol{A}_1\boldsymbol{x}_{t-1} + \ldots + \boldsymbol{A}_p\boldsymbol{x}_{t-p} + \boldsymbol{u}_t$$

where

- $\boldsymbol{a}_0$ is a $m-vector$ of intercepts,
- $\boldsymbol{A}_k, k = 1,\ldots,p$ are square $m \times m$ matrices,
- $\boldsymbol{u}_t$ $(m \times 1)$ is a vector of disturbances with

$$E(\boldsymbol{u}_t) = \mathbf{0}$$

$$E(\boldsymbol{u}_t\boldsymbol{u}_t') = \boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & \cdots & \omega_{1m} \\ \vdots & \ddots & \vdots \\ \omega_{m1} & \cdots & \omega_{mm} \end{bmatrix}$$

Also write

$$\boldsymbol{A}(L)\boldsymbol{x}_t = \boldsymbol{u}_t$$

where $\boldsymbol{A}(L) = \boldsymbol{I} - \boldsymbol{A}_1L - \cdots - \boldsymbol{A}_pL^p$.

# Stability of the VAR

Start with the case $p = 1$.

Solve the system by repeated substitution as

$$\begin{aligned} \boldsymbol{x}_t &= \boldsymbol{a}_0 + \boldsymbol{A}\boldsymbol{x}_{t-1} + \boldsymbol{u}_t \\ &= \boldsymbol{a}_0 + \boldsymbol{A}(\boldsymbol{a}_0 + \boldsymbol{A}\boldsymbol{x}_{t-2} + \boldsymbol{u}_{t-1}) + \boldsymbol{u}_t \\ &= \cdots \\ &= (\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2 + \ldots)\boldsymbol{a}_0 + \boldsymbol{u}_t + \boldsymbol{A}\boldsymbol{u}_{t-1} + \boldsymbol{A}^2\boldsymbol{u}_{t-2} + \cdots \end{aligned}$$

Thus, we need to know the properties of $\boldsymbol{A}^n$ as $n$ gets large .

# Eigenvalues

The eigenvalues of $\boldsymbol{A}$ are the solutions (assumed distinct - possibly complex-valued) to the equation

$$|\boldsymbol{A} - \mu\boldsymbol{I}| = \boldsymbol{0}.$$

A square matrix with distinct eigenvalues $\mu_1, \mu_2, \ldots, \mu_m$ has diagonalization $\boldsymbol{A} = \boldsymbol{CMC}^{-1}$ where $\boldsymbol{M} = \mathrm{diag}(\mu_1, \ldots, \mu_m)$

and $\boldsymbol{C}$ is the matrix of eigenvectors.

Note that

$$\begin{aligned} |\boldsymbol{A} - \mu\boldsymbol{I}| &= |\boldsymbol{CMC}^{-1} - \mu\boldsymbol{I}| = |\boldsymbol{C}(\boldsymbol{M} - \mu\boldsymbol{C}^{-1}\boldsymbol{C})\boldsymbol{C}^{-1}| = |\boldsymbol{M} - \mu\boldsymbol{I}||\boldsymbol{C}||\boldsymbol{C}^{-1}| = |\boldsymbol{M} - \mu\boldsymbol{I}| \\ &= (\mu_1 - \mu)(\mu_2 - \mu)\cdots(\mu_m - \mu). \end{aligned}$$

Note

$$\begin{aligned} \boldsymbol{A}^n &= \boldsymbol{CMC}^{-1}\boldsymbol{CMC}^{-1}\cdots\boldsymbol{CMC}^{-1} \\ &= \boldsymbol{CM}^n\boldsymbol{C}^{-1}. \end{aligned}$$

The conditions for $\boldsymbol{A}^n \to \boldsymbol{0}$ as $n \to \infty$ are that $|\mu_i| < 1$, for each $i$.

# The General Case

Write the VAR($p$) model in *companion form:*

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{x}_{t-1} \\ \boldsymbol{x}_{t-2} \\ \vdots \\ \boldsymbol{x}_{t-p} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_p & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{t-1} \\ \boldsymbol{x}_{t-2} \\ \vdots \\ \boldsymbol{x}_{t-p} \\ \boldsymbol{x}_{t-p-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{u}_t \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

or

$$\boldsymbol{x}_t^* = \boldsymbol{A}^* \boldsymbol{x}_{t-1}^* + \boldsymbol{u}_t^* \qquad (m(p+1) \times 1).$$

- Repeat the analysis of $p = 1$ on this model.
- $\boldsymbol{A}^*$ $(m(p+1) \times m(p+1))$ has $m(p+1)$ eigenvalues of which $m$ are 0.
- Can be shown that the remaining $mp$ eigenvalues match the inverted roots of $|\boldsymbol{A}(\lambda)| = 0$.

# The Generalized Stability (Invertibility) Condition:

Generalizing the AR($p$) analysis, we can show that all the roots of the equation

$$|\boldsymbol{A}(\lambda)| = |\lambda^p \boldsymbol{I} - \lambda^{p-1}\boldsymbol{A}_1 - \cdots - \boldsymbol{A}_p|$$

(a polynomial of order $mp$) must lie outside the unit circle.

- Note the case $p = 1$. Observe that

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \lambda|\boldsymbol{I} - \mu \boldsymbol{A}|$$

  where $\mu = 1/\lambda$, so the eigenvalue and polynomial root conditions are equivalent.

- Note the case $m = 1$. The companion form provides an alternative way to analyse the stability of the AR($p$).

# The Final Form of a VAR

To solve $\boldsymbol{A}(L)\boldsymbol{x}_t = \boldsymbol{u}_t$, note that

$$\boldsymbol{A}(L)^{-1} = \frac{1}{|\boldsymbol{A}(L)|}\,\mathrm{adj}\,\boldsymbol{A}(L)$$

where $|\boldsymbol{A}(\lambda)|$ is a lag polynomial of order $mp$ and the elements of $\mathrm{adj}\,\boldsymbol{A}(L)$ are lag polynomials of maximum order $p(m-1)$.

Hence the final form equations are

$$|\boldsymbol{A}(L)|\boldsymbol{x}_t = \mathrm{adj}\,\boldsymbol{A}(L)\boldsymbol{u}_t$$

Key facts:

- The vector on the right-hand side is a sum of $m$ moving average terms in the elements of $\boldsymbol{u}_t$.
- Can be shown that a sum of moving averages has a MA representation.

Hence, the final form of a VAR is a vector of univariate ARMA processes!

- Nominally the AR roots are the same for each equation.
- In practice there are frequently cancellations of common factors, element by element.