HG
Jan 2014 ECON 5101

## Exercises II - 10 Feb 2014

**Exercise 1.**
Read section 8 in lecture notes 3 (LN3) on the common factor problem in ARMA-processes.

Consider the following process

(1) $Y_t = 0.4Y_{t-1} + 0.45Y_{t-2} + \varepsilon_t + \varepsilon_{t-1} + 0.25\varepsilon_{t-2}$

where $\varepsilon_t \sim WN(0,\sigma^2)$.

This is an ARMA process of the form $\varphi(L)Y_t = \theta(L)\varepsilon_t$ without constant term.

**A.** Investigate if the two lag-polynomials have a common factor. If so, reformulate the difference equation specification for $\{Y_t\}$ to a proper ARMA specification.

**B.** Is there a causal stationary solution for $\{Y_t\}$? If yes, write up the solution as a $MA(\infty)$ process. Is the MA specification invertible? If so, write up the $AR(\infty)$ solution for $\varepsilon_t$ .

**Exercise 2.**
**A.** Suppose that $Y_t \sim ARMA(p,q)$ and causal, satisfying ) $\varphi(L)Y_t = \varphi_0 + \theta(L)\varepsilon_t$, where $\varphi(L) = 1 - \varphi_1 L - \cdots - \varphi_p L^p$, $\theta(L) = (1 + \theta_1 L + \cdots + \theta_q L^q)$. Introduce the centered series, $y_t = Y_t - \mu$, leading to $\varphi(L)y_t = \theta(L)\varepsilon_t$ without the constant $\varphi_0$, and where $E(y_t) = 0$ (see exercise 3 of seminar1 ).

Explain why statement (7) on page 5 of LN3, saying
$\varphi(L)\gamma(h) = 0$ and $\varphi(L)\rho(h) = 0$ for $h \geq \max(p, q+1)$, is true, where $\gamma(h), \rho(h)$ are the autocovariance function and acf respectively.

[**Hint.** Assume $h \geq \max(p, q+1)$ and write

$$\gamma(h) = E(y_{t+h}y_t) = E\left(\varphi_1 y_{t+h-1}y_t + \cdots + \varphi_p y_{t+h-p}y_t + \varepsilon_{t+h}y_t + \cdots + \theta_q \varepsilon_{t+h-q}y_t\right)$$

etc. ]

B. **Introduction**. If $Y_t \sim AR(1)$ with $y_t = \varphi y_{t-1} + \varepsilon_t$ and $|\varphi| < 1$, we get from page 3 in LN2 that the autocovariance function, $\gamma_1(-h) = \gamma_1(h) = \sigma^2 \dfrac{\varphi^h}{1-\varphi^2}$ which implies that the acf is $\rho(-h) = \rho(h) = \varphi^h$ for $h = 0,1,2,\ldots$

We are interested in this exercise to find out what the effect is on the AR(1) acf, $\rho(h)$ by adding a MA-term, $\theta\varepsilon_{t-1}$ to $y_t = \varphi y_{t-1} + \varepsilon_t$.

Assume therefore that $y_t - \varphi y_{t-1} = \varepsilon_t + \theta\varepsilon_{t-1}$, i.e., an ARMA(1,1), where $|\varphi| < 1, \quad |\theta| < 1,$ and $\varepsilon_t \sim WN(0,\sigma^2)$. We need the autocovariance, $\gamma(h)$ and acf $\rho(h)$ for $y_t$ in this new situation. First $\gamma(h)$: Using (8) on page 5 in LN3 and **A.**, we have

(i) $\gamma(h) - \varphi\gamma(h-1) = 0$ for $h = 2,3,\cdots$

with initial conditions

(ii) $\gamma(0) - \varphi\gamma(-1) = \sigma^2(\delta_0 + \theta\delta_1)$ , i.e., $\gamma(0) - \varphi\gamma(1) = \sigma^2(\delta_0 + \theta\delta_1)$

(iii) $\gamma(1) - \varphi\gamma(0) = \sigma^2(\theta\delta_0)$

where $\delta_0, \delta_1$ are found from $\dfrac{1+\theta z}{1-\varphi z} = \delta_0 + \delta_1 z + \delta_2 z^2 + \cdots$

**Question.** Show that $\delta_0 = 1$ and $\delta_1 = \varphi + \theta$

[**Hint.** Write the lag-equation (in terms of companion series) $1 + \theta z = (1-\varphi z)(\delta_0 + \delta_1 z + \delta_2 z^2 + \cdots)$ . Then multiply out the right side sufficiently to determine $\delta_0, \delta_1$ . ]

**C.** Show first from ii. and iii. that

$$\gamma(0) = \sigma^2 \frac{1 + 2\theta\varphi + \theta^2}{1-\varphi^2}$$

$$\gamma(1) = \sigma^2 \frac{(\theta+\varphi)(1+\theta\varphi)}{1-\varphi^2}$$

and then that $\gamma(h) = \sigma^2 \dfrac{(\theta+\varphi)(1+\theta\varphi)}{1-\varphi^2}\varphi^{h-1}$ for $h \geq 1$ .

**D.** Show that the acf of $y_t$ can be expressed as $\rho(h) = \dfrac{(\theta+\varphi)(1+\theta\varphi)}{\varphi(1+2\theta\varphi+\theta^2)}\rho_1(h)$,

where $\rho_1(h)$ is the acf of the AR(1) process.

**E.** Looking at the constant factor in front of $\rho_1(h)$, characterize the effect $\theta$ has on the AR(1) acf in the two special cases, $\varphi = 0.9$ and $\varphi = -0.2$ – i.e., for which values of $\theta$ will the AR(1) acf increase and for which values will it decrease?

**[Hint.** The factor in front of $\rho_1(h)$ is not so easy to discuss analytically. The best way, in my view, is simply to plot the factor as a function of $\theta$ for $-1 < \theta < 1$ in the two cases. The easiest is probably to plot it with a computer. This is, for example, easy in Stata. The following command, for example,

```
twoway function y=2*x^2-1, range(-1 2.5)
```

plots the function $y = 2x^2 - 1$ for $-1 < x < 2.5$ ]

**F.** For any $\varphi$ with $|\varphi| < 1$, is there any value of $\theta$ that turns $\{y_t\}$ into white noise? If so, for which value?

## Appendix Some basic mathematical facts about power series that lies under the theory of finding solutions of stochastic difference equations.

This (maybe too long) appendix is by no means meant to be compulsory reading (and not as necessary for solving the exercises!) but more as an attempt to explain for the curious student some of the math that is really lying underneath much of our manipulations in the lectures. If it contributes to the understanding for some, it is fine. If not, it may not matter so much. Working with concrete examples and basic concepts may be just as efficient in creating good understanding over time… Note that I in this appendix do not go into elementary Hilbert space theory which is also needed. Maybe later.

Sometimes a function $f(z)$ can be expressed as an infinite sum of power terms in *z*,

(A1) $$f(z) = \delta_0 + \delta_1 z + \delta_2 z^2 + \cdots + \delta_n z^n + \cdots$$

Such a series is called *a power series*. It satisfies very strong mathematical properties that underlie important applied branches of mathematics - including the theory of dynamic stochastic models that involves stochastic difference equations. A famous example of a power series that you probably have seen, is the exponential function

(A2) $$e^z = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \cdots + \frac{1}{n!}z^n + \cdots$$

This function is well defined for any real (or complex) *z*. Its *radius of convergence* (see fact 1) is $\infty$.

*I will describe some facts that you should know.*

**Fact 1.** The function, $f(z)$, in (A1) is well defined and convergent (i.e., "absolutely convergent" to use a technical term) everywhere inside a certain open symmetrical interval, $(-r, r)$ about zero (i.e., for $|z| < r$). This is in case *z* is a real variable. If *z* is a complex variable, it is also well defined. It actually turns out to be well defined everywhere inside a certain circle defined by $|z| < r$. The border of convergence, *r*, is called *the radius of convergence*. On the border, $|z| = r$, anything can happen, it sometimes converges and sometimes not. Notice in particular that the constant term is always given by $\delta_0 = f(0)$.

> [**Comment.** I will not go into proofs of this here since you apparently have learned very little (if not nothing) about convergent series (infinite sums) in your math training. We won't need that theory here, but I may, for completeness sake, just mention how convergence of series are studied in math. One basically looks at truncated versions of (1), i.e, $s_n(z) = \delta_0 + \delta_1 z + \delta_2 z^2 + \cdots + \delta_n z^n$ - which are just finite polynomials - and study the conditions for convergence of the *sequence*, $\{s_n(z)\}_{n=0}^{\infty}$ of such functions.

Another thing worth mentioning, is the "*root – criterion*" (one of several criteria for convergence of infinite series) for determining the radius of convergence in (1), namely[1],

(A3) $$r=\frac{1}{\lim\limits_{n\to\infty}\left(\sqrt[n]{|\delta_n|}\right)}=\frac{1}{\lim\limits_{n\to\infty}\left(|\delta_n|^{\frac{1}{n}}\right)}$$

For example, you know about geometric series resulting in the expansion

(A4) $$\begin{aligned}\frac{1}{1-\varphi z}&=1+\varphi z+(\varphi z)^2+(\varphi z)^3+\cdots\\&=1+\varphi z+\varphi^2 z^2+\varphi^3 z^3+\cdots\end{aligned}$$

which is an example of a power series. In this series, $\delta_n=\varphi^n$, and the root criterion determines *r* as

(A5) $$r=\frac{1}{\lim\limits_{n\to\infty}\left(\sqrt[n]{|\delta_n|}\right)}=\frac{1}{\lim\limits_{n\to\infty}\left(\sqrt[n]{|\varphi|^n}\right)}=r=\frac{1}{\lim\limits_{n\to\infty}\left(|\varphi|\right)}=\frac{1}{|\varphi|}$$

We see that if $|\varphi|<1$, then $r>1$, showing that the expansion (A4) is well defined in the open interval $(-r,r)$ which includes $z=\pm 1$ as inner points (or, in the case of complex *z*, includes the whole unit circle as inner points). This implies among other things that $\frac{1}{1-\varphi}=f(1)$ is well defined.

There are many ways of generating power series. For example, Taylor expansion around zero generates a power series within the circle of convergence (or interval in the real case). **End of comment.** ]

**Fact 2.** The following is a beautiful theorem on power series that is the maybe the main cause of many major developments in mathematics:

**i.** Inside the circle (interval in the real case) of convergence ($|z|<r$ ), the function, $f(z)$, in (1) is continuously differentiable and the derivative can be found by derivating the series term by term causing a new power series

---

[1] or, more generally, $r=\frac{1}{\lim\sup\limits_{n\to\infty}\left(\sqrt[n]{|\delta_n|}\right)}$, in case the limit in the denominator does not exist.

(A6) $$f'(z) = \delta_1 + 2\delta_2 z + 3\delta_3 z^2 + \cdots + n\delta_n z^{n-1} + \cdots$$

**ii.** The power series for $f'(z)$ in (A6) has the same (!) radius of convergence, *r*, as the series for $f(z)$.

[**Comment.** The proof of **ii.** follows directly from the root criterion in (A5), considering the "well known" result (in math books) that the sequence $\sqrt[n]{n}$ converges to 1 as $n \to \infty$. The proof of **i.** is a little tricky, but is not long, and can be found in most introductory text books on complex analysis.

The theorem now implies that also $f'(z)$ is continuously differentiable within the circle/interval of convergence with derivative (having the same radius of convergence)

$$f''(z) = 2\delta_2 + 3\cdot 2\delta_3 z + \cdots + n(n-1)\delta_n z^{n-2} + \cdots$$

and so on…

An interesting conclusion is that an expansion like (1) of $f(z)$ with radius of convergence, *r*, implies that $f(z)$ is infinitely many times differentiable for all *z* with $|z| < r$, and the *n*-th derivative, $f^{(n)}(z)$ can be found by derivating the power series term by term *n* times. Doing that, note that the constant term in the series for $f^{(n)}(z)$ then must be

(A7) $$f^{(n)}(0) = n!\delta_n$$

**End of comment**. ]

From **fact 2** we can now conclude the (for us) important fact 3.

**Fact 3.** The coefficients $\{\delta_j\}$ of the power series expansion of $f(z)$ in (1) are uniquely determined. In other words: If

$$f(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \cdots = \beta_0 + \beta_1 z + \beta_2 z^2 + \cdots$$

are two power series expansions for $f(z)$, all corresponding coefficients must be equal (i.e., $\alpha_j = \beta_j$ for $j = 0,1,2,\ldots$.

[**Proof.** We must have $f(0)=\alpha_0=\beta_0$ , $f'(0)=\alpha_1=\beta_1$ , … ,
$f^{(n)}(0)=n!\alpha_n=n!\beta_n$ implying $\alpha_n=\beta_n$ for all *n*. **End of proof**.

**Example 3.1:** The constant function $f(z)=1$ for all *z*, has a power series expansion (trivially)

$$f(z)=1+0z+0z^2+0z^3+\cdots$$

The uniqueness of the expansion tells us that there is no other way to express 1 (or any other constant) in a power series.

**Example 3.2**: I have used this uniqueness property several times in the lectures without saying so. On some occasions we needed to solve an equation

$$\varphi(z)\psi(z)=1$$

with respect to the power series, $\psi(z)=\psi_0+\psi_1 z+\psi_2 z^2+\cdots$, where $\varphi(z)=1-\varphi_1 z-\cdots-\varphi_p(z)$ is a polynomial (i.e., a special kind of power series with all except a finite number of coefficients = 0). The solution for $\psi(z)$ is found by multiplying out the two power series on the left (see fact 4 below) and identify the result with the power series on the right (i.e., the constant 1). The uniqueness then implies that all coefficients you find on the left side – except the constant term – must be equal to the corresponding coefficients on the right, namely 0.
**End of comment**. ]

Now we can establish a full algebra of power series enabling us to define new power series from old ones by addition and multiplication. The rules are given in fact 4.

**Fact 4.**

**i.** Multiplying the power series in (1) by any constant $c\neq 0$ gives a new power series with the same radius of convergence as $f(z)$,

$$cf(z)=c\delta_0+c\delta_1 z+c\delta_2 z^2+\cdots+c\delta_n z^n+\cdots$$

[The radius of convergence must be the same as before by the root criterion since $\sqrt[n]{|c|}\underset{n\to\infty}{\longrightarrow}1$ ]

**ii.** Let $f_1(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \cdots$ and $f_2(z) = \beta_0 + \beta_1 z + \beta_2 z^2 + \cdots$ be two power series and *r* the smallest of the two radiuses of convergence involved. Then $f_1(z) + f_2(z)$ is a power series with radius of convergence *r* and where the coefficients are obtained by termwise addition

$$f_1(z) + f_2(z) = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)z + (\alpha_2 + \beta_2)z^2 + \cdots$$

**iii.** Also the product $f_1(z)\, f_2(z)$ is a power series with radius of convergence *r*, and where the coefficients are obtained by usual multiplication of terms and collecting terms with the same power of $z$ (see comment).

[**Comment**. The formal proof is relatively simple (although a little messy), but requires some theory from general convergence of series and is therefore skipped. The process of multiplication goes along in the same way as multiplication of polynomials where we multiply one of the polynomial systematically by every term in the other and add up – as illustrated in the following table

| Factor | Result |
|---|---|
| $\alpha_0$ | $\alpha_0\beta_0 + \alpha_0\beta_1 z + \alpha_0\beta_2 z^2 + \cdots + \alpha_0\beta_n z^n + \ldots$ |
| $\alpha_1 z$ | $\alpha_1\beta_0 z + \alpha_1\beta_1 z^2 + \alpha_1\beta_2 z^3 + \cdots + \alpha_1\beta_n z^{n+1} + \cdots$ |
| $\alpha_2 z^2$ | $\alpha_2\beta_0 z^2 + \alpha_2\beta_1 z^3 + \alpha_2\beta_2 z^4 + \cdots + \alpha_2\beta_n z^{n+2} + \cdots$ |
| $\vdots$ | $\vdots$ |
| $\alpha_n z^n$ | $\alpha_n\beta_0 z^n + \alpha_n\beta_1 z^{n+1} + \alpha_n\beta_2 z^{n+2} + \cdots + \alpha_n\beta_n z^{n+n} + \cdots$ |
| $\vdots$ | $\vdots$ |

Summing up and collecting terms we get

$$f_1(z)\, f_2(z) = \delta_0 + \delta_1 z + \delta_2 z^2 + \cdots + \delta_n z^n + \cdots =$$
$$= \alpha_0\beta_0 + (\alpha_0\beta_1 + \alpha_1\beta_0)z + (\alpha_0\beta_2 + \alpha_1\beta_1 + \alpha_2\beta_0)z^2 + \cdots$$

So, by uniqueness,

$$\delta_0 = \alpha_0\beta_0, \quad \delta_1 = \alpha_0\beta_1 + \alpha_1\beta_0, \quad \ldots$$
$$\delta_n = \alpha_0\beta_n + \alpha_1\beta_{n-1} + \alpha_2\beta_{n-2} + \cdots + \alpha_{n-1}\beta_1 + \alpha_n\beta_0 \;\; nnn$$

**End of comment.** ]

## Application to linear filters used on causal stationary time series.

Now, let $\{y_t\}_{t=-\infty}^{\infty}$ be any causal stationary time series. Note that we assume that the series has been going on since $-\infty$. This is just a trick. In reality the (observed) series has started, of course, at some fixed time point, say $t=0$. The variables $y_{-1}, y_{-2}, y_{-3}, \ldots$ are just imagined and not observed (i.e., *latent*) variables with the same statistical properties as the observed ones, $y_0, y_1, y_2, \ldots$ The trick enables us in a simple way to construct causal stationary solutions (that are stationary right from the start $t=0$) for a large class of stochastic difference equation by using infinite linear filters constructed from power series:

**Fact 5.**
If $f(z)=\delta_0+\delta_1 z+\delta_2 z^2+\cdots+\delta_n z^n+\cdots$ is *any* power series with radius of convergence $r>1$ (this is essential), then it defines uniquely a linear filter

(A8) $$f(L) \overset{DEF}{=} \delta_0+\delta_1 L+\delta_2 L^2+\cdots+\delta_n L^n+\cdots$$

that operates on $\{y_t\}_{t=-\infty}^{\infty}$ by

(A9) $$w_t = f(L)\,\mathrm{y}_t \overset{DEF}{=} \delta_0 y_t+\delta_1 y_{t-1}+\delta_2 y_{t-2}+\cdots+\delta_n y_{t-n}+\cdots$$

The corresponding power series function, $f(z)$, is sometimes called *the companion power series*. In this way we have established a one-to-one mathematical relationship between power series (with $r>1$ ) and infinite linear filters operating on causal stationary series. Such a filter satisfies the following properties (this theorem 2 in LN2):

**i.** $w_t = f(L)y_t$ defined in (A9) is a well-defined random variable and the series $\{w_t\}$ is covariance stationary with expectation $E(w_t)=f(L)\mu=f(1)\mu$, where $\mu=E(y_t)$, and a certai autocovariance function (not given here except in the case $y_t$ is white noise (see (5) in LN2 page 2).

[**Note.** It is here we need some elementary Hilbert space theory (not discussed here)for a formal proof.]

**ii.** The filter $f(L)$ is linear in the sense that if $\{y_t\},\{z_t\}$ are two stationary series, then $f(L)\left(ay_t+bz_t\right)=af(L)y_t+bf(L)z_t$, where *a,b* are constants.

**iii.** If $f_1(L), f_2(L)$ are two such linear filters (with $r > 1$), then the product $f(L) = f_1(L)f_2(L)$ also is such a filter (with $r > 1$) and with coefficients determined by multiplication of the two companion power series.

**iv.** An alternative (to convergence radius > 1) condition that essentially gives the same conclusions about $f(L)$, is the one cited in the lectures, i.e., the condition that

$$\sum_{j=0}^{\infty} |\delta_j| < \infty .$$

[**Comment**. A complete formal proof is not given here. As an example, suppose $y_t$ is expressed by two other stationary series $\{x_t\}, \{z_t\}$ and white noise error $\varepsilon_t$, as $y_t = c + \theta_1(L)x_t + \theta_2(L)z_t + \theta_3(L)\varepsilon_t$ where $c$ is a constant and $\theta_j(L), j = 1,2,3$ are three proper filters. Suppose we want to use $f(L)$ on $y_t$, we get using ii. and iii., that

$$w_t = f(L)\mathrm{y}_t = f(1)c + f(L)\theta_1(L)x_t + f(L)\theta_2(L)z_t + f(L)\theta_3(L)\varepsilon_t$$

showing how the influences the explanatory stationary series in the equation.
**End of comment**. ]