STAT 804: Notes on Lecture 3

Defn: If $\{\epsilon_t\}$ is a white noise series and μ and $b_0, \ldots, b_p$ are constants then

$$X_t = \mu + b_0\epsilon_t + b_1\epsilon_{t-1} + \cdots + b_p\epsilon_{t-p}$$

is a moving average of order p; write $MA(p)$.

Q: From observations on X can we estimate the b's and $\sigma^2 = \text{Var}(\epsilon_t)$ accurately? NO.

Defn: **Model** for data X is family $\{P_\theta; \theta \in \Theta\}$ of possible distributions for X.

Defn: Model is **identifiable** if $\theta_1 \neq \theta_2$ implies $P_{\theta_1} \neq P_{\theta_2}$; different θ's give different distributions for data.

Unidentifiable model: there are different values of θ which make exactly the same predictions about the data.

So: data do not distinguish between these θ values.

Example: Suppose ϵ is an iid $N(0, \sigma^2)$ series and that $X_t = b_0\epsilon_t + b_1\epsilon_{t-1}$. Then the series X has mean 0 and covariance

$$C_X(h) = \begin{cases} (b_0^2 + b_1^2)\sigma^2 & h = 0 \\ b_0 b_1 \sigma^2 & h = 1 \\ 0 & \text{otherwise} \end{cases}$$

Fact: normal distribution is specified by its mean and its variance.

Consequence: two mean 0 normal time series and the same covariance function have the same distribution.

Observe: if you multiply the ϵ's by a and divide both b_0 and b_1 by a then the covariance function of X is unchanged.

Thus: cannot hope to estimate all three parameters, b_0, b_1 and σ.

Arbitrary choice: $b_0 = 1$

Are parameters b_1 and σ identifiable?

We try to solve the equations

$$C(0) = (1 + b^2)\sigma^2$$

and

$$C(1) = b\sigma^2$$

to see if the solution is unique. Divide the two equations to see

$$\frac{C(1)}{C(0)} = \frac{b}{1 + b^2}$$

or

$$b^2 - \frac{C(0)}{C(1)} b + 1 = 0$$

which has the solutions

$$\frac{\frac{C(0)}{C(1)} \pm \sqrt{\left(\frac{C(0)}{C(1)}\right)^2 - 4}}{2}$$

You should notice two things:

1. If
$$\left|\frac{C(0)}{C(1)}\right| < 2$$
there are no solutions.

   Since $C(0) = \sqrt{\text{Var}(X_t)\text{Var}(X_{t+1})}$ we see $C(1)/C(0)$ is the correlation between X_t and X_{t+1}.

   So: have proved that for an $MA(1)$ process this correlation cannot be more than 1/2 in absolute value.

2. If
$$\left|\frac{C(0)}{C(1)}\right| > 2$$
there are two solutions.

Note: two solutions multiply together to give the constant term 1 in the quadratic equation.

If two roots are distinct it follows that one of them is larger than 1 and the other smaller in absolute value.

Let b and b^* denote the two roots.

Let $\alpha = C(1)/b$ and $\alpha^* = C(1)/b^*$.

Let ϵ_t be iid $N(0, \alpha)$ and ϵ_t^* be iid $N(0, \alpha^*)$. Then

$$X_t \equiv \epsilon_t + b\epsilon_{t-1}$$

and

$$X_t^* \equiv \epsilon_t^* + b^*\epsilon_{t-1}^*$$

have identical means and covariance functions. Observing X_t you cannot distinguish the first of these models from the second. We will fit $MA(1)$ models by **requiring** our estimated b to have $|\hat{b}| \leq 1$.

Reason: manipulate model equation for X as for autoregressive process:

$$\begin{aligned}\epsilon_t &= X_t - b\epsilon_{t-1} \\ &= X_t - b(X_{t-1} - b\epsilon_{t-2}) \\ &\vdots \\ &= \sum_0^\infty (-b)^j X_{t-j}\end{aligned}$$

This manipulation makes sense if $|b| < 1$. If so then we can rearrange the equation to get

$$X_t = \epsilon_t - \sum_1^\infty (-b)^j X_{t-j}$$

which is an autoregressive process.

If, on the other hand, $|b| > 1$ then we can write

$$X_t = \frac{1}{b} b\epsilon_t - b\epsilon_{t-1}$$

Let $\epsilon_t^* = b\epsilon_t$; ϵ^* is also white noise. We find

$$\begin{aligned}
\epsilon_{t-1}^* &= X_t - \frac{1}{b}\epsilon_t^* \\
&= X_t - \frac{1}{b}(X_{t+1} - \frac{1}{b}\epsilon_{t+1}^*) \\
&\vdots \\
&= \sum_0^\infty (-\frac{1}{b})^j X_{t+j}
\end{aligned}$$

which means

$$X_t = \epsilon_{t-1}^* - \sum_1^\infty (-\frac{1}{b})^j X_{t+j}$$

This represents the current value as depending on the future which seems physically far less natural than the other choice.

Defn: An $MA(p)$ process is invertible if it can be written in the form

$$X_t = \sum_1^\infty a_j X_{t-j} + \epsilon_t$$

Defn: A process X is an autoregression of order p (written $AR(p)$) if

$$X_t = \sum_1^p a_j X_{t-j} + \epsilon_t$$

(so an invertible MA is an infinite order autoregression).

Defn: The backshift operator transforms a time series into another time series by shifting it back one time unit; if X is a time series then BX is the time series with

$$(BX)_t = X_{t-1}\,.$$

The identity operator I satisfies $IX = X$. We use B^j for $j = 1, 2, \ldots$ to denote B composed with itself j times so that

$$(B^j X)_t = X_{t-j}$$

For $j = 0$ this gives $B^0 = I$.

Now use B to develop a formal method for studying the existence of a given $AR(p)$ and the invertibility of a given $MA(p)$.

An $AR(1)$ process satisfies

$$(I - a_1 B)X = \epsilon$$

Think of $I - a_1 B$ as infinite dimensional matrix; get formal identity

$$X = (I - a_1 B)^{-1}\epsilon$$

So how will we define this inverse of an infinite matrix? We use the idea of a geometric series expansion.

If b is a real number then

$$(1 - ab)^{-1} = \frac{1}{1 - ab} = \sum_{j=0}^{\infty} (ab)^j$$

so we hope that $(I - a_1 B)^{-1}$ can be defined by

$$(I - a_1 B)^{-1} = \sum_{j=0}^{\infty} a_1^j B^j$$

This would mean

$$X = \sum_{j=0}^{\infty} a_1^j B^j \epsilon$$

or looking at the formula for a particular t and remembering the meaning of B^j we get

$$X_t = \sum_{j=0}^{\infty} a_1^j \epsilon_{t-j}$$

This is the formula I had in lecture 2.

Now consider a general $AR(p)$ process:

$$(I - \sum_1^p a_j B^j) X = \epsilon$$

We will factor the operator applied to x. Let

$$\phi(x) = 1 - \sum_1^p a_j x^j$$

Then ϕ is degree p polynomial so it has (theorem of C. F. Gauss) p roots $1/b_1, \ldots, 1/b_p$. (None of the roots is 0 because the constant term in ϕ is 1.) This means we can factor ϕ as

$$\phi(x) = \prod_1^p (1 - b_j x)$$

Now back to the definition of X:

$$\prod_1^p (I - b_j B) X = \epsilon$$

can be solved by inverting each term in the product (in any order — the terms in the product commute) to get

$$X = \prod_1^p (I - b_j B)^{-1} \epsilon$$

The inverse of $I - b_1 B$ will exist if the sum

$$\sum_{k=0}^{\infty} b_j^k B^k$$

converges; this requires $|b_j| < 1$. Thus a stationary $AR(p)$ solution of the equations exists if every root of the characteristic polynomial ϕ is larger than 1 in absolute value (actually the roots can be complex and I mean larger than 1 in modulus).

Summary

- An $MA(q)$ process $X_t = \epsilon_t - \sum_{j=1}^q b_j \epsilon_{t-j}$ is invertible iff all roots of characteristic polynomial $\psi(x) = 1 - \sum_{j=1}^q b_j x^j$ lie outside unit circle in complex plain.

- For given covariance function of an $MA(q)$ process there is only one set of coefficients $b_1, \ldots, b_q$ for which the process is invertible.

- An $AR(p)$ process $X_t - \sum_{j=1}^q a_j X_{t-j} = \epsilon_t$ is asymptotically stationary iff all roots of characteristic polynomial $\phi(x) = 1 - \sum_1^p a_j x^j$ lie outside unit circle in complex plain.

  (Asymptotically stationary means: make $X_{-1}, X_{-2}, \ldots, X_{-p}$ anything; use equation defining $AR(p)$ to define rest of X values; then as $t \to \infty$ the process gets closer to being stationary.

  Asymptotic stationarity is equivalent to existence of an exactly stationary solution of equations.)

Defn: A process X is an $ARMA(p,q)$ (mixed autoregressive of order p and moving average of order q) if it satisfies

$$\phi(B)X = \psi(B)\epsilon$$

where ϵ is white noise and

$$\phi(B) = I - \sum_1^p a_j B^j$$

and

$$\psi(B) = I - \sum_1^p b_j B^j$$

The ideas we used above can be stretched to show that the process X is identifiable and causal (can be written as an infinite order autoregression on the past) if the roots of $\psi(x)$ lie outside the unit circle. A stationary solution, which can be written as an infinite order causal (no future ϵs in the average) moving average, exists if all the roots of $\phi(x)$ lie outside the unit circle.

Other Stationary Processes:

Periodic processes: Suppose Z_1 and Z_2 are independent $N(0, \sigma^2)$ random variables and that ω is a constant. Then

$$X_t = Z_1 \cos(\omega t) + Z_2 \sin(\omega t)$$

has mean 0 and

$$\begin{aligned} \mathrm{Cov}(X_t, X_{t+h}) &= \sigma^2 [\cos(\omega t)\cos(\omega(t+h)) \\ &\quad + \sin(\omega t)\sin(\omega(t+h))] \\ &= \sigma^2 \cos(\omega h) \end{aligned}$$

Since X is Gaussian we find that X is second order and strictly stationary. In fact (see your homework) You can write

$$X_t = R \sin(\omega t + \Phi)$$

where R and Φ are suitable random variables so that the trajectory of X is just a sine wave.

Poisson shot noise processes:

Poisson process is a process $N(A)$ indexed by subsets A of $\mathbb{R}$ such that each $N(A)$ has a Poisson distribution with parameter λlength(A) and if $A_1, \ldots A_p$ are any non-overlapping subsets of R then $N(A_1), \ldots, N(A_p)$ are independent. We often use $N(t)$ for $N([0,t])$.

Shot noise process: $X(t) = 1$ at those t where there is a jump in N and 0 elsewhere; X is stationary.

If g a function defined on $[0, \infty)$ and decreasing sufficiently quickly to 0 (like say $g(x) = e^{-x}$) then the process

$$Y(t) = \sum g(t-\tau) 1(X(\tau) = 1) 1(\tau \le t)$$

is stationary.

Y jumps every time t passes a jump in Poisson process; otherwise follows trajectory of sum of several copies of g (shifted around in time). We commonly write

$$Y(t) = \int_0^\infty g(t-\tau) dN(\tau)$$