STAT 804: Notes on Lecture 5

Model identification

By **model identification** for a time series X we mean the process of selecting values of p, q so that the $ARMA(p,q)$ process gives a reasonable fit to our data. The most important model identification tool is a plot of (an estimate of) the autocorrelation function of X; we use the abbreviation ACF for this function. Before we discuss doing this with real data we explore what plots of the ACF of various $ARMA(p,q)$ plots should look like (in the absence of estimation error).

For an $MA(p)$ process we found that

$$C_X(h) = \begin{cases} \sigma^2 \sum_{j=0}^{p-|h|} b_j b_{j+|h|} & |h| \le p \\ 0 & \text{otherwise} \end{cases}$$

This has the important *qualitative* feature that it vanishes for $|h| > p$.

For an $AR(1)$ process $X_t - \mu = \rho(X_{t-1} - \mu) + \epsilon_t$ the autocorrelation function is

$$\rho_X(h) = \rho^{|h|}$$

which has the qualitative feature of decreasing geometrically.

To derive the autocovariance for a general $AR(p)$ we mimic the technique for $p = 1$. If $X_t = \sum_1^p a_j X_{t-j} + \epsilon_t$ then

$$\begin{aligned} C_X(h) &= \text{Cov}(X_0, X_h) \\ &= \sum_{j=1}^{p} a_j \text{Cov}(X_0, X_{h-j}) + \text{Cov}(X_0, \epsilon_h) \\ &= \sum_{j=1}^{p} a_j C_X(h-j) \end{aligned}$$

for $h > 0$. Take these equations and divide through by $C_X(0)$ and remember that $\rho_X(h) = C_X(h)/C_X(0)$ and $\rho_X(-k) = \rho_X(k)$ you see that the above recursions for $h = 1, \ldots, p$ are p linear equations in the p unknowns $\rho_X(1), \ldots, \rho_X(p)$. They are called the Yule Walker equations. For instance, when $p = 2$ we get

$$\begin{aligned} C_X(2) &= a_1 C_X(1) + a_2 C_X(0) \\ C_X(1) &= a_1 C_X(0) + a_2 C_X(-1) \end{aligned}$$

which becomes, after division by $C_X(0)$

$$\rho_X(2) = a_1\rho_X(1) + a_2$$
$$\rho_X(1) = a_1 + a_2\rho_X(1)$$

It is possible to use generating functions to get explicit formulas for the $\rho(h)$ but here we simply observe that we have two equations in two unknowns to solve. The second equation shows that

$$\rho(1) = \frac{a_1}{1-a_2}$$

which is not possible if $a_2 = 1$ (unless $a_1 = 0$) and not a correlation for some other (a_1, a_2) pairs. The first equation then gives

$$\rho(2) = \frac{a_1^2 + a_2(1-a_2)}{1-a_2}$$

Notice that the Yule Walker equations permit $\rho(h)$ to be calculated recursively from $\rho(1)$ and $\rho(2)$ for $h \geq 3$.

Now look at $\phi(x)$, the characteristic polynomial, when $a_2 = 1$ we have

$$\phi(x) = 1 - a_1 x - x^2 = (1-\alpha_1 x)(1-\alpha_2 x)$$

where $1/\alpha_i, i = 1, 2$ are the two roots. Multiplying out we find that $\alpha_1\alpha_2 = -1$ so that either one of the two has modulus more than 1 (and the root $1/\alpha_i$ has modulus less than 1) or both have modulus 1. The two roots may be seen to be real so they would have to be ± 1. Since $\alpha_1 + \alpha_2 = a_1$ (again from multiplying it out and examining the coefficient of x) we would then know $a_1 = 0$. In either case there is no stationary solution.

Qualitative features: It is possible to prove that the solutions of these Yule-Walker equations decay to 0 at a geometric rate meaning that they satisfy $|\rho_X(h)| \leq a^{|h|}$ for some $a \in (0, 1)$. However, for general p they are not too simple.

Periodic Processes

If Z_1, Z_2 are iid $N(0, \sigma^2)$ then we saw

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

is a strictly stationary process with mean 0 and autocorrelation$\rho(h) = \cos(\omega h)$. Thus the autocorrelation would be perfectly periodic.

Linear Superposition

If X and Y are jointly stationary then $Z = aX + bY$ is stationary and

$$C_Z(h) = a^2 C_X(h) + b^2 C_Y(h) + ab(C_{XY}(h) + C_{YX}(h))$$

Thus you could hope, for example, to recognize a periodic component to a series by looking for a periodic component to a plotted autocorrelation.

Periodic versus AR processes

In fact you can make AR processes which behave very much like periodic processes. Consider the process

$$X_t = X_{t-1} - aX_{t-2} + \epsilon_t$$

Here are graphs of trajectories and autocorrelations for $a = 0.3, 0.6, 0.9$ and 0.99.

You should observe the slow decay of the waves in the autocovariances, particularly for a near 1. When $a = 1$ the characteristic polynomial is $1 - x + x^2$ which has roots

$$\frac{1 \pm \sqrt{-3}}{2}$$

Both these roots have modulus 1 so there is no stationary trajectory with $a = 1$. The point is that some AR processes have nearly periodic components.

To get more insight consider the differential equation describing a sine wave:

$$\frac{d^2}{dx^2} f(x) = -\omega^2 f(x);$$

the solution if $f(x) = a\sin(\omega x + \phi)$. If we replace the derivative by differences we get the approximation

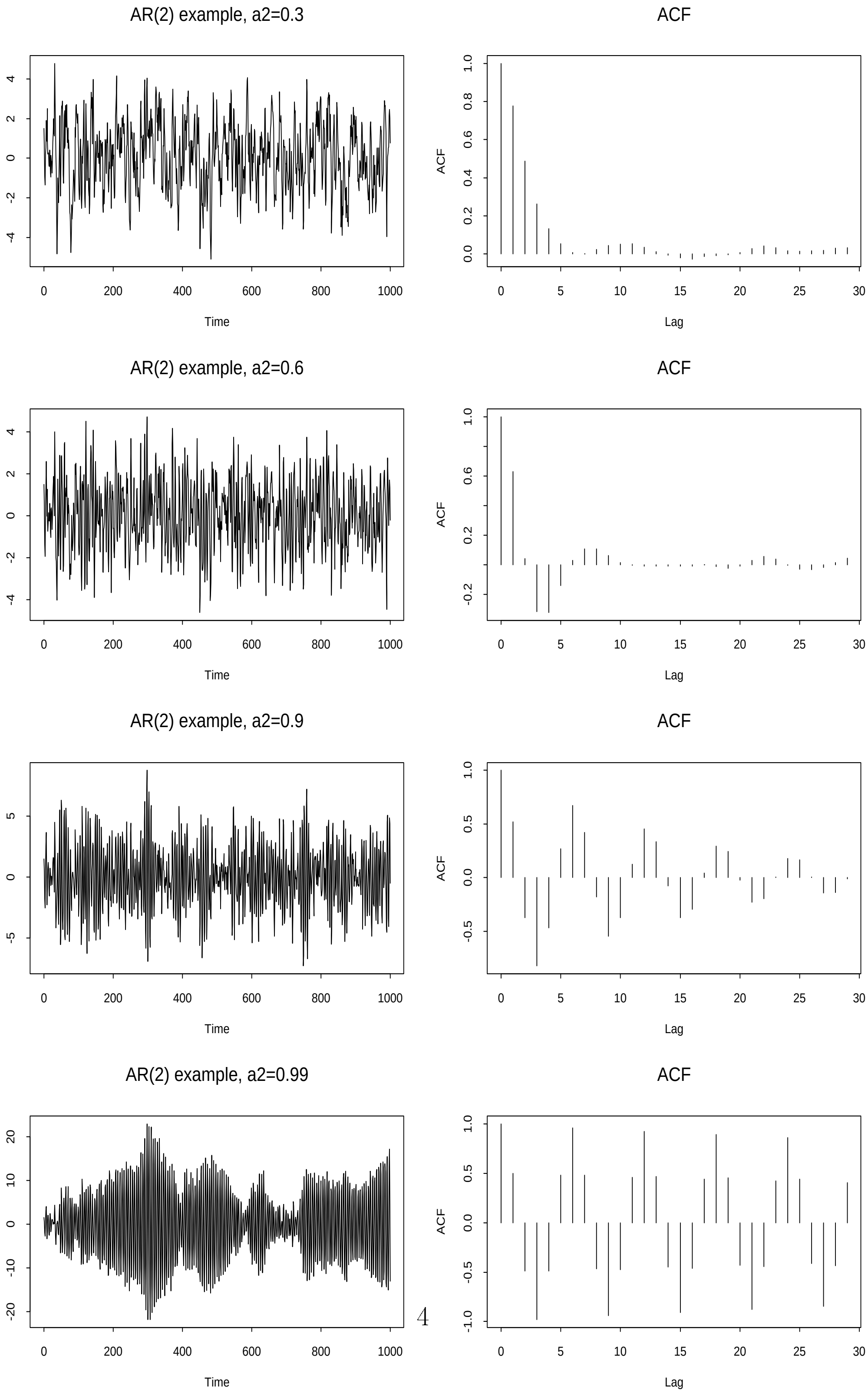
$$\frac{d^2}{dx^2} f(x) \approx \frac{f(x+h) - 2f(x) + f(x-h}{h^2}$$

so that

$$\frac{f(x+h) - 2f(x) + f(x-h}{h^2} \approx -\omega^2 f(x)$$

Take $h = 1$ in the approximation and reorganize to get

$$f(x+1) = (2 - \omega^2) f(x) - f(x-1)$$


AR(2) example, a2=0.3
ACF
Time
Lag
AR(2) example, a2=0.6
ACF
Time
Lag
AR(2) example, a2=0.9
ACF
Time
Lag
AR(2) example, a2=0.99
ACF
Time
Lag

If we add noise, change notation to $t = x + 1$ and replace the letter f by X we get

$$X_t = (2 - \omega^2) X_{t-1} - X_{t-2} + \epsilon_t$$

This is formalism only; there is no stationary solution of this equation. However, we see that $AR(2)$ processes are at least analogous to the solutions of second order differential equations with added noise.

Estimates of C and ρ

In order to identify suitable $ARMA$ models using data we need estimates of C and ρ. If we knew that $\mu = 0$ we would see that

$$C_X(h) = \text{Cov}(X_0, X_h) = \text{Cov}(X_1, X_{h+1}) = \cdots$$

We would then be motivated to use

$$\hat{C}(h) = \sum_{0}^{T-1-h} X_t X_{t+h} / T$$

simply averaging products over all pairs which are h time units apart. When μ is unknown we will often simply use $\hat{\mu} = \bar{X}$ and then take

$$\hat{C}(h) = \sum_{0}^{T-1-h} (X_t - \hat{\mu})(X_{t+h} - \hat{\mu}) / T$$

or, noting that there are only $T - h$ terms in the sum

$$\hat{C}(h) = \sum_{0}^{T-1-h} (X_t - \hat{\mu})(X_{t+h} - \hat{\mu}) / (T - h)$$

We then take

$$\hat{\rho}(h) = \hat{C}(h) \hat{C}(0)$$

(Note, however, that when $T - h$ is used in the divisor it is technically possible to get a $\hat{\rho}$ value which exceeds 1.)