

# STAT 804: Lecture 5

## Model identification

**Model identification** for time series  $X$ : select values of  $p, q$  so that the  $ARMA(p, q)$  process gives a reasonable fit to data.

Most important tool: plot of estimated autocorrelation function (ACF) of  $X$ .

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

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

Important *qualitative* feature: vanishes if  $|h| > p$ .

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

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

qualitative feature: decreases geometrically.

To derive ACF for 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  $h > 0$ .

Divide through by  $C_X(0)$ .

Remember that  $\rho_X(h) = C_X(h)/C_X(0)$  and  $\rho_X(-k) = \rho_X(k)$ : see that the above recursions for  $h = 1, \dots, p$  are  $p$  linear equations in the  $p$  unknowns  $\rho_X(1), \dots, \rho_X(p)$ .

Called the Yule Walker equations.

For instance, when  $p = 2$  we get

$$C_X(2) = a_1 C_X(1) + a_2 C_X(0)$$

$$C_X(1) = a_1 C_X(0) + a_2 C_X(-1)$$

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

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

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

Can use generating functions to get explicit formulas for the  $\rho(h)$ .

Here simply observe: two equations in two unknowns to solve.

The second equation shows that

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

Not possible if  $a_2 = 1$  (unless  $a_1 = 0$ )

Not a correlation for some other  $(a_1, a_2)$  pairs (see homework).

The first equation then gives

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

Note: can calculate  $\rho(h)$  recursively from  $\rho(1)$  and  $\rho(2)$  for  $h \geq 3$  via Yule Walker.

Look at characteristic polynomial  $\phi(x)$ :

When  $a_2 = 1$  we have

$$\phi(x) = 1 - a_1x - x^2 = (1 - \alpha_1x)(1 - \alpha_2x)$$

where  $1/\alpha_i, i = 1, 2$  are the two roots.

Multiplying out:  $\alpha_1\alpha_2 = -1$  so either:

One of two has modulus more than 1 (root  $1/\alpha_i$  has modulus less than 1) or

Both have modulus 1.

Both roots real so would be  $\pm 1$ .

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:** can prove solutions of Yule-Walker equations decay to 0 at a geometric rate:  $|\rho_X(h)| \leq a^{|h|}$  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 strictly stationary, mean 0, autocorrelation  $\rho(h) = \cos(\omega h)$ : perfectly periodic.

## Linear Superposition

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

$$\begin{aligned} C_Z(h) &= a^2 C_X(h) + b^2 C_Y(h) \\ &\quad + ab(C_{XY}(h) + C_{YX}(h)) \end{aligned}$$

Could hope to recognize periodic component in series by looking for a periodic component in plotted ACF.

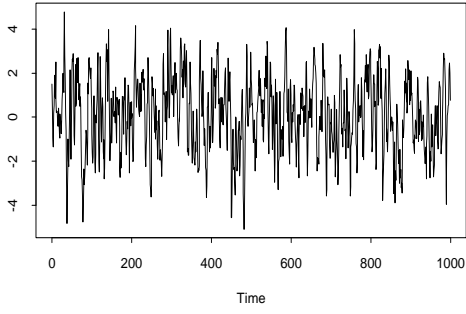
## Periodic versus AR processes

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

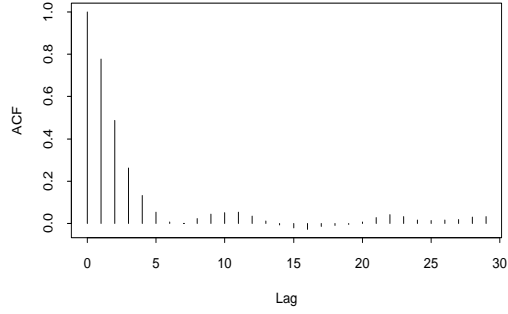
$$X_t = X_{t-1} - aX_{t-2} + \epsilon_t$$

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

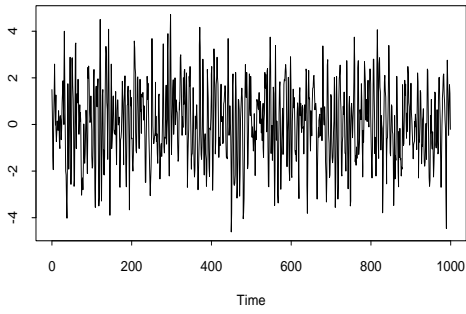
AR(2) example,  $a_2=0.3$



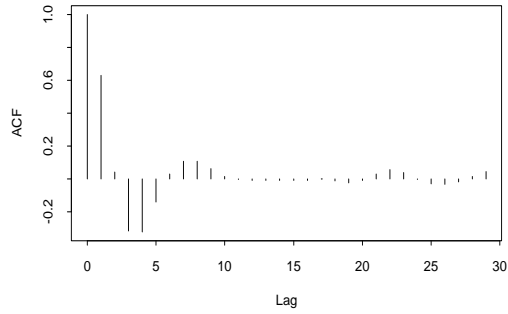
ACF



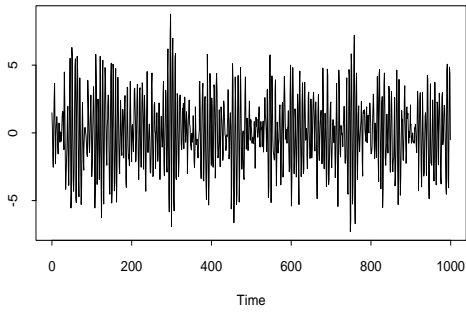
AR(2) example,  $a_2=0.6$



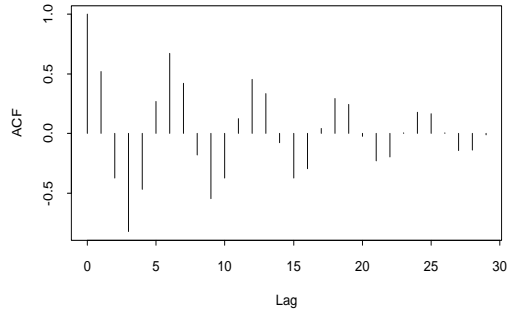
ACF



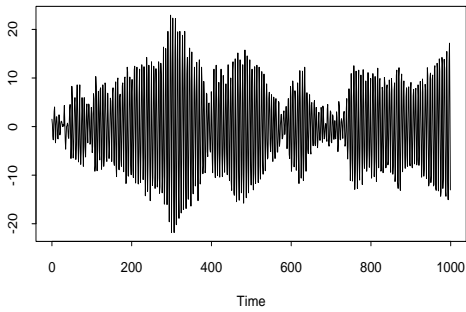
AR(2) example,  $a_2=0.9$



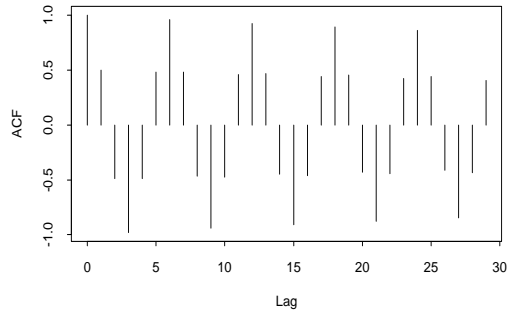
ACF



AR(2) example,  $a_2=0.99$



ACF



Observe: slow decay of waves in autocovariances, particularly for  $a$  near 1.

When  $a = 1$  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:

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

solution is  $f(x) = a \sin(\omega x + \phi)$ . Replace derivative by differences: get 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  $h = 1$  and reorganize to get

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

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

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

Formalism only; no stationary solution exists.

But,  $AR(2)$  processes are at least analogous to solutions of second order differential equations with added noise.

### **Estimates of $C$ and $\rho$**

In order to identify suitable  $ARMA$  models using data we need estimates of  $C$  and  $\rho$ . If  $\mu = 0$  is known then

$$\begin{aligned} C_X(h) &= \text{Cov}(X_0, X_h) = \text{Cov}(X_1, X_{h+1}) = \dots \\ &= \text{E}(X_0 X_h) = \text{E}(X_1 X_{h+1}) = \dots \end{aligned}$$

We would then be motivated to use

$$\hat{C}(h) = \sum_0^{T-1-h} X_t X_{t+h} / T.$$

Average products over all pairs which are  $h$  time units apart. When  $\mu$  is unknown often simply use  $\hat{\mu} = \bar{X}$ ; take

$$\hat{C}(h) = \sum_0^{T-1-h} (X_t - \hat{\mu})(X_{t+h} - \hat{\mu}) / T$$

Alternative: only  $T - h$  terms in the sum

$$\hat{C}(h) = \sum_0^{T-1-h} (X_t - \hat{\mu})(X_{t+h} - \hat{\mu}) / (T - h).$$

so use

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

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