Econometrics 2b: Handout #4

Matthew J. Lindquist*

March 2009

1 The Autocorrelation Function

- **Assuming weak stationarity**, we can define the k-th order autocovariance, γ_k, as
$$\gamma_k = cov\{Y_t, Y_{t-k}\} = cov\{Y_t, Y_{t+k}\}.$$

- The autocovariance of a stochastic process can be standardized and presented as an autocorrelation function (ACF), ρ_k
$$\rho_k = \frac{cov\{Y_t, Y_{t-k}\}}{V\{Y_t\}} = \frac{\gamma_k}{\gamma_0}.$$
The ACF helps characterize the development of Y_t over time. It shows us how strongly current observations are correlated with past observations and how shocks today affect future values of the stochastic variable. Besides helping us describe the data, the ACF will also help us find unit roots, choose models and run regression diagnostics.

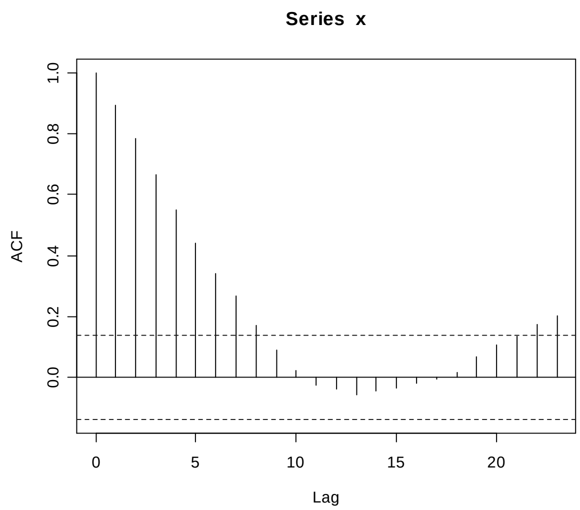
- The ACF of an $AR(1)$ process is
$$\rho_k = \frac{cov\{Y_t, Y_{t-k}\}}{V\{Y_t\}} = \frac{\theta^{|k-t|}\frac{\sigma^2}{1-\theta^2}}{\frac{\sigma^2}{1-\theta^2}} = \theta^k.$$

- The ACF of an $MA(1)$ process is
$$\rho_k = \frac{cov\{Y_t, Y_{t-k}\}}{V\{Y_t\}} = \frac{\alpha\sigma^2}{(1+\alpha^2)\sigma^2} = \frac{\alpha}{1+\alpha^2}.$$

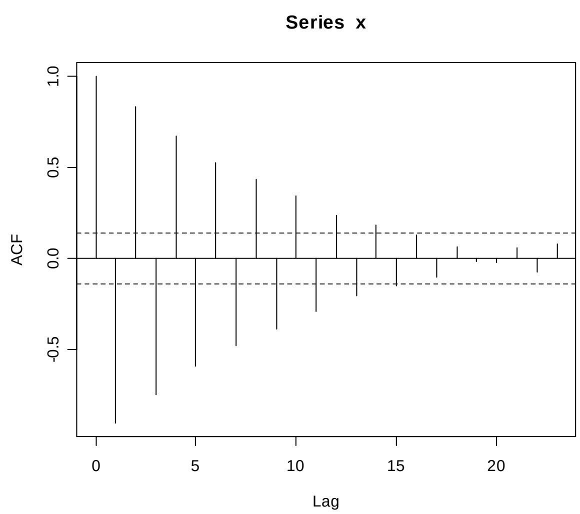
- In $\mathbb{R}$ you can graph the ACF of a variable, x, by writing: **acf(x)**

*Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden. Tel. 46+8+163831. E-mail: Matthew.Lindquist@ne.su.se.

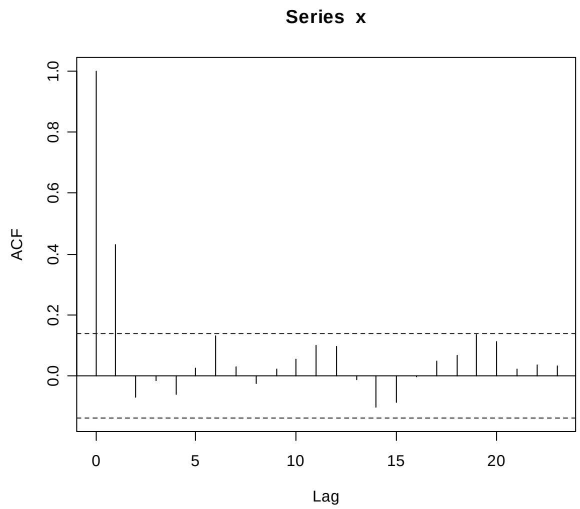
```
x <- arima.sim(list(order=c(0,0,1), ar=0.90), n=200)
acf(x)
```



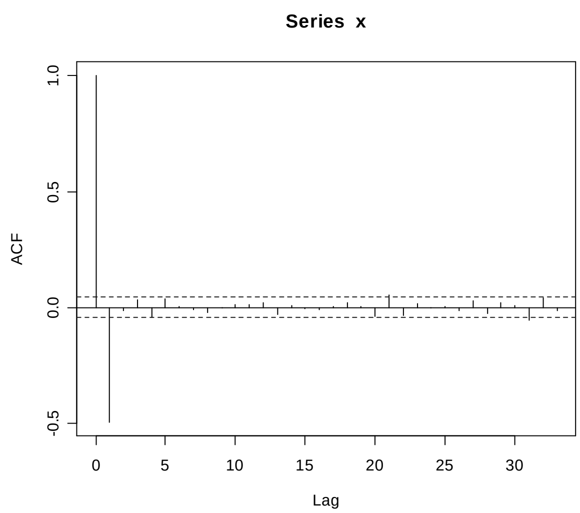
The ACF of an AR(1) With $\theta = 0.90$.



The ACF of an AR(1) Process With $\theta = -0.90$.



The ACF of a MA(1) process with $\alpha = 0.90$.



The ACF of a MA(1) Process With $\alpha = -0.90$.

- It is often times difficult to distinguish between different AR(q) processes based solely on an examination of a correlogram.

2 The Partial Autocorrelation Function

- The partial autocorrelation function (PACF) may provide a more clear discrimination.

- The δ_2 parameter in an AR(2) process is the partial correlation coefficient between x_t and x_{t-2} holding x_{t-1} constant.

- Recall the definition of a partial correlation coefficient in the 3 variable case.

$$\delta_2 = r_{13.2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{23}^2}}.$$

- **Assuming weak stationarity**

$$r_{12} = corr\left\{x_t, x_{t-1}\right\} = corr\left\{x_{t-1}, x_{t-2}\right\} = \rho_1$$

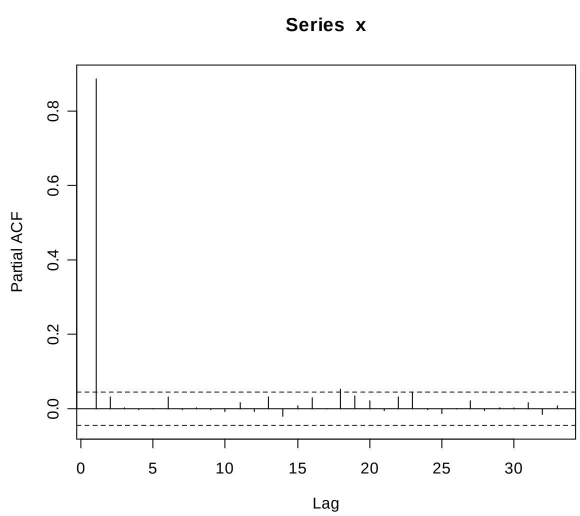
  and

$$r_{13} = corr\left\{x_t, x_{t-2}\right\} = \rho_2.$$

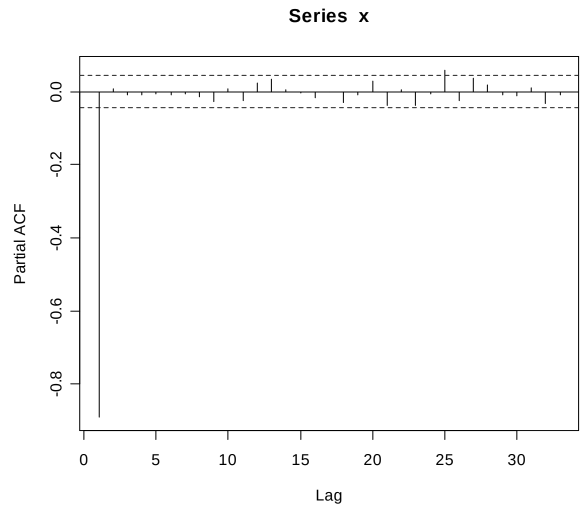
  Substituting this into the formula for δ_2

$$\delta_2 = r_{13.2} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}.$$

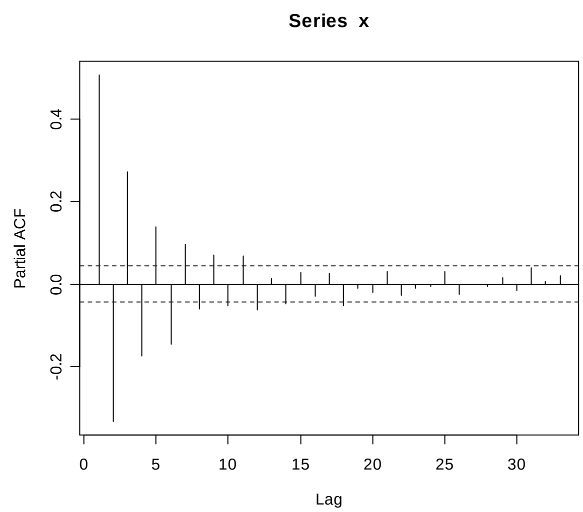
- These can be graphed in $\mathbb{R}$ by writing: **pacf(x)**



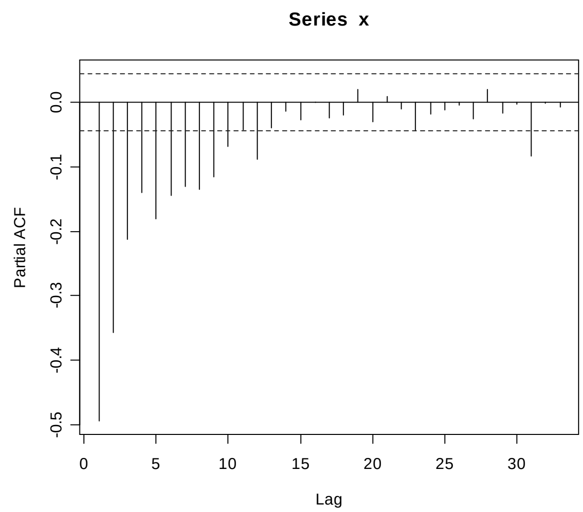
The PACF of an AR(1) Process With $\theta = 0.90$.



The PACF of an AR(1) Process With $\theta = -0.90$.



The PACF of a MA(1) Process With $\alpha = 0.90$.



The PACF of a MA(1) Process With $\alpha = -0.90$.

3 Summary of Results

- An $AR(p)$ process is described by:

  1. an ACF that is infinite in extent (it tails off)
  2. a PACF that is (close to) zero for lags larger than p.

- A $MA(q)$ process is described by:

  1. an ACF that is (close to) zero for lags larger than q.
  2. a PACF that is infinite in extent (it tails off).

- Note that this "identification" strategy only works on stationary time series! For example, if you have quarterly data that trends upwards, you must first remove the seasonal pattern and then remove the long-run trend. After this is done, you can use the acf and pacf functions to identify the de-trended variable.