UCI Paul Merage School of Business

Leadership for a Digitally Driven World™

MFIN 290: Financial Econometrics

Lecture 4-2

This Time

- Stationarity Example
- Cointegration and Error Correction

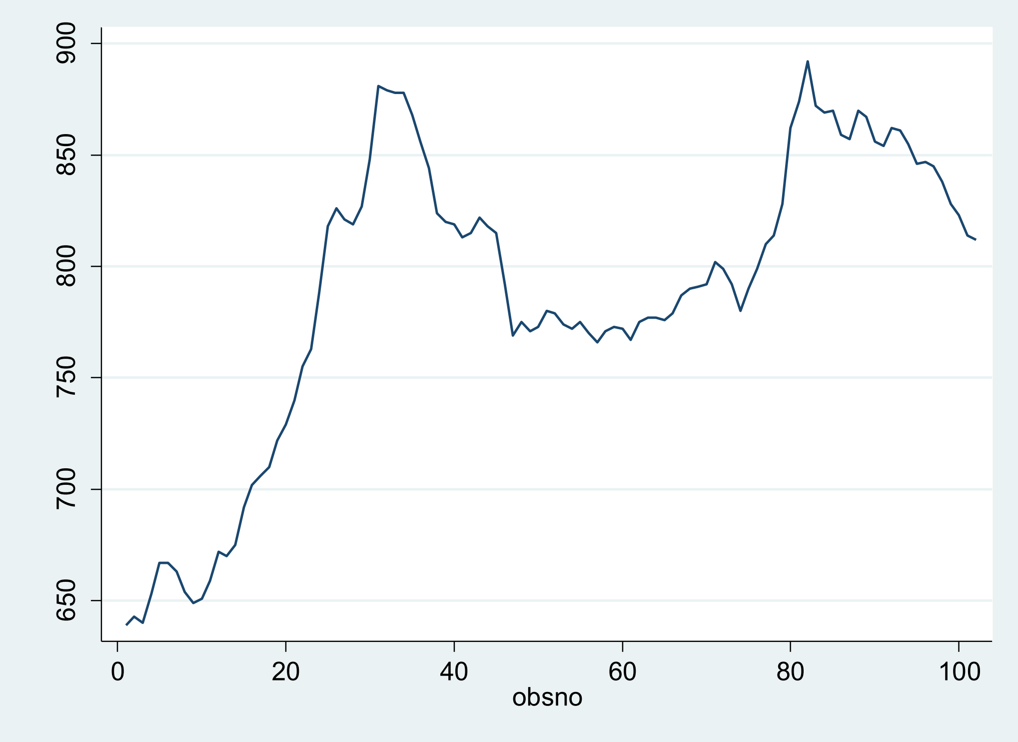
- Yule Walker Equations
- ARMA(1,1)

Stationarity Example

- AA Railroad bond yields
- January 1968 – June 1976
- Bond.dta

Example

```
. gen obsno = _n

. tsset obsno
        time variable:  obsno, 1 to 102
                delta:  1 unit

. tsline bond
```



Example

- Fail to reject. Remember the null is that the series is non-stationary!

```
. dfuller bond

Dickey-Fuller test for unit root                   Number of obs   =       101

                              ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.361            -3.510            -2.890            -2.580
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.1530
```

Let's try differencing

Example

Can have it show the full regression as well….

```
. gen bond_diff = bond - l.bond
(1 missing value generated)
```

```
. dfuller bond_diff, regress

Dickey-Fuller test for unit root                   Number of obs   =       100

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -5.955            -3.510            -2.890            -2.580
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

------------------------------------------------------------------------------
 D.bond_diff |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   bond_diff |
         L1. |  -.5318553    .089308    -5.96   0.000    -.7090842   -.3546265
             |
       _cons |   .8707468   .9412997     0.93   0.357    -.9972319    2.738725
------------------------------------------------------------------------------
```

Order of Integration

Series are sometimes talked about in terms of their "order of integration".

This is the number of times it must be differenced before becoming stationary and is written I(N).

Stationary series? I(0)
Random Walk? I(1)

Our railroad bond example appears stationary after a single differencing => it is I(1)

Order of Integration

Note again that differencing series will only make them stationary with random walks.

Autocorrelated series will not suddenly lose their autocorrelation after differencing (though the autocovariances will change!)

KPSS Test

Problem with ADF (well, one problem), is that we have a null hypothesis of non-stationarity. **We can never accept the null with any test**… so depending on what you are trying to conclude, you may not be able to do it with that test.

KPSS test looks for stationarity using a null that the series is stationary (with more traditional test stats as well).

Non-Stationary Data

In the last lecture, we talked about why we wanted to work with stationary data to ensure we avoid spurious regression problems.

Non-Stationary Data

However, there are specific circumstances where not only is it acceptable to work with non-stationary series, it may be preferred!

If y_t and x_t are both non-stationary variables, with the same order of integration, we ought to expect the difference between y_t and x_t – or any linear combination of them – to have the same order of integration.

Non-Stationary Data

However, there are specific circumstances where not only is it acceptable to work with non-stationary series, it may be preferred!

If y_t and x_t are both non-stationary variables we ought to expect the difference between y_t and x_t – or any linear combination of them – to still be non-stationary.

However, when there is a linear function:

$$y_t = \beta_0 + \beta_1 x_t + e_t$$

That yields a *stationary series* e_t, more is possible….

Cointegration

$$y_t = \beta_0 + \beta_1 x_t + e_t$$

If e_t is stationary, then it means that x_t and y_t never move too far from one another (they must be drifting together at some common rate).

Motivation: consider the following system where y_1 and y_2 trend through time:

$$y_{1t} = \alpha + \beta t + u_t$$
$$y_{2t} = \gamma + \delta t + v_t$$

Where u_t, v_t are white noise.

Cointegration

$$y_{1t} = \alpha + \beta t + u_t$$
$$y_{2t} = \gamma + \delta t + v_t$$

Consider the linear combination $y_{1t} + \theta y_{2t}$

$$z_t = y_{1t} + \theta y_{2t} = \alpha + \theta\gamma + (\beta + \theta\delta)t + u_t + \theta v_t$$

Cointegration

$$y_{1t} = \alpha + \beta t + u_t$$
$$y_{2t} = \gamma + \delta t + v_t$$

Consider the linear combination $y_{1t} + \theta y_{2t}$

$$z_t = y_{1t} + \theta y_{2t} = \alpha + \theta\gamma + (\beta + \theta\delta)t + u_t + \theta v_t$$

Which will generally still be I(1), unless $(\beta + \theta\delta) = 0$ …

$(\beta + \theta\delta) = 0 \Rightarrow \theta = -\frac{\beta}{\delta}$

Cointegration

$$y_{1t} = \alpha + \beta t + u_t$$
$$y_{2t} = \gamma + \delta t + v_t$$

That is, if we can take a linear combination of the variables and remove the common time trend, we will be left with an I(0) (stationary) series after that particular linear combination.

Cointegration

$$y_{1t} = \alpha + \beta t + u_t$$
$$y_{2t} = \gamma + \delta t + v_t$$

These linear combinations are referred to as "cointegrating vectors". Here, the cointegrating vector would be

$y_{1t} + \theta y_{2t} = [1, \theta] = \left[1, -\frac{\beta}{\delta}\right]$ to get a stationary series of residuals.

This turns out to be unique – the ONLY way that two series can be cointegrated is if they share a common trend.

Related: Error-Correction models

Can also have relationships between I(0) series that embed cointegrating relationships.

Imagine z_t and y_t are I(1) series, and are cointegrated with cointegrating vector $(1, -\theta)$.

=> $y_t - y_{t-1}, z_t - z_{t-1},$ and $y_t - \theta z_t$ are all I(0) series.

Error-Correction models

Consider the model

$$\Delta y_t = x_t\beta + \gamma\Delta \mathrm{z_t} + \lambda(y_{t-1} - \theta z_{t-1}) + e_t$$

The changes in y around some long term trend are a function of some exogenous factors x_t, how z varies around its long term trend, and the error correction $(y_{t-1} - \theta z_{t-1})$, which is the equilibrium error from the model of cointegration before.

Note: If y_t and z_t are both non-stationary, but not co-integrated with cointegrating vector $[1, -\theta]$, then $y_{t-1} - \theta z_{t-1}$ cannot be I(0).

Error-Correction models

$$\Delta y_t = x_t\beta + \gamma\Delta z_t + \lambda(y_{t-1} - \theta z_{t-1}) + e_t$$

This suggests that the errors from a cointegrating model can inform the behavior of the levels of the series – if the cointegrated series drift too far apart, they will be "pulled" back together.

If we know θ, we can estimate this equation, should get the same betas, but now everything in this equation is stationary, so we will get the right inference

This is a big advantage.

Properties of Cointegrated Series

The Paul Merage School of Business

Normally, we could not estimate an equation like $y_t = \beta_0 + \beta_1 x_t + e_t$

If y_t and x_t were not covariance stationary.

Not only do we worry about spurious regression, but in principle, y_t and x_t both have their own time series equations with their own trend, and they need to be determined simultaneously.

Properties of Cointegrated Series

Cointegrated series do not have this problem! The estimator is what is called "superconsistent": it converges to the true values *more rapidly* than an OLS estimator.

$\frac{1}{T^2}$ vs. $\frac{1}{T}$ in the OLS case

But the t-statistics are not interpretable because the variance is not bounded.

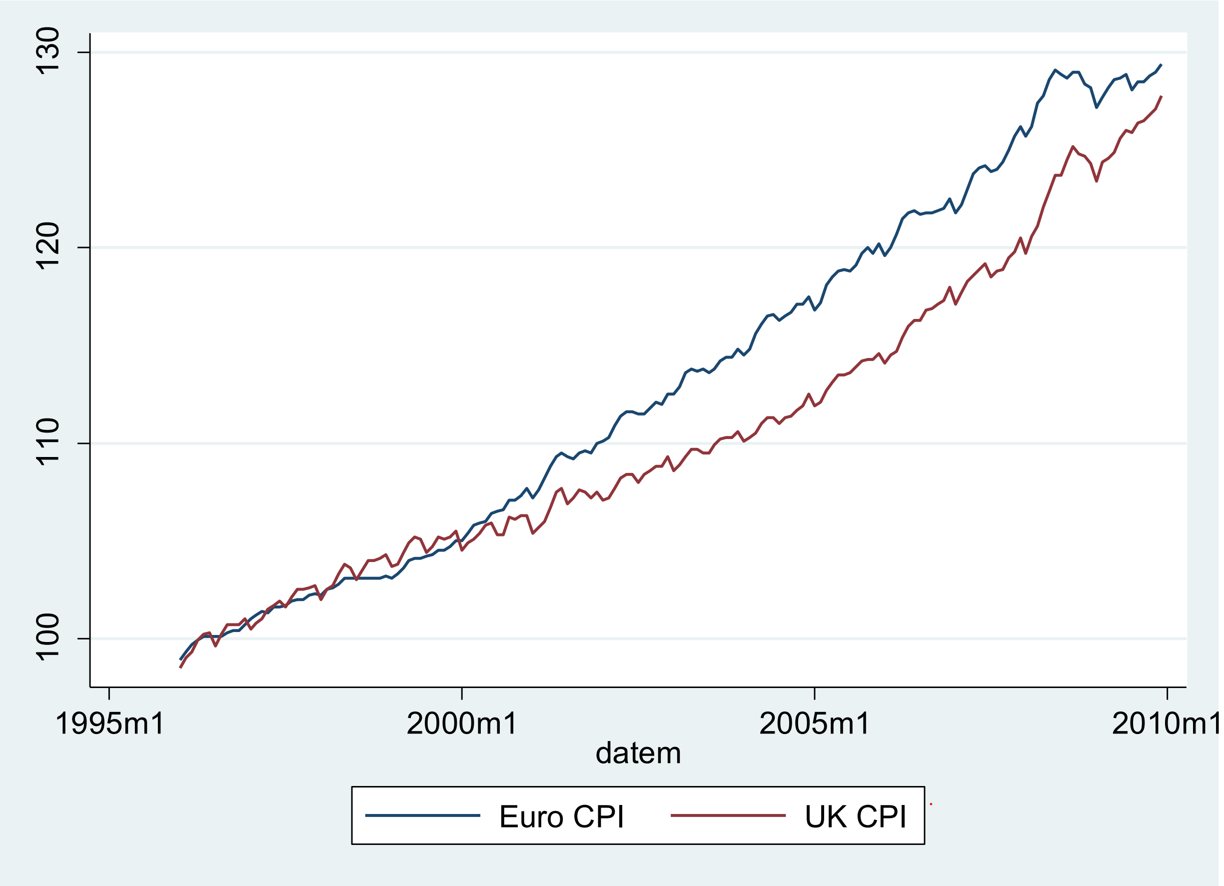
Cointegration Example

ukpi.dta

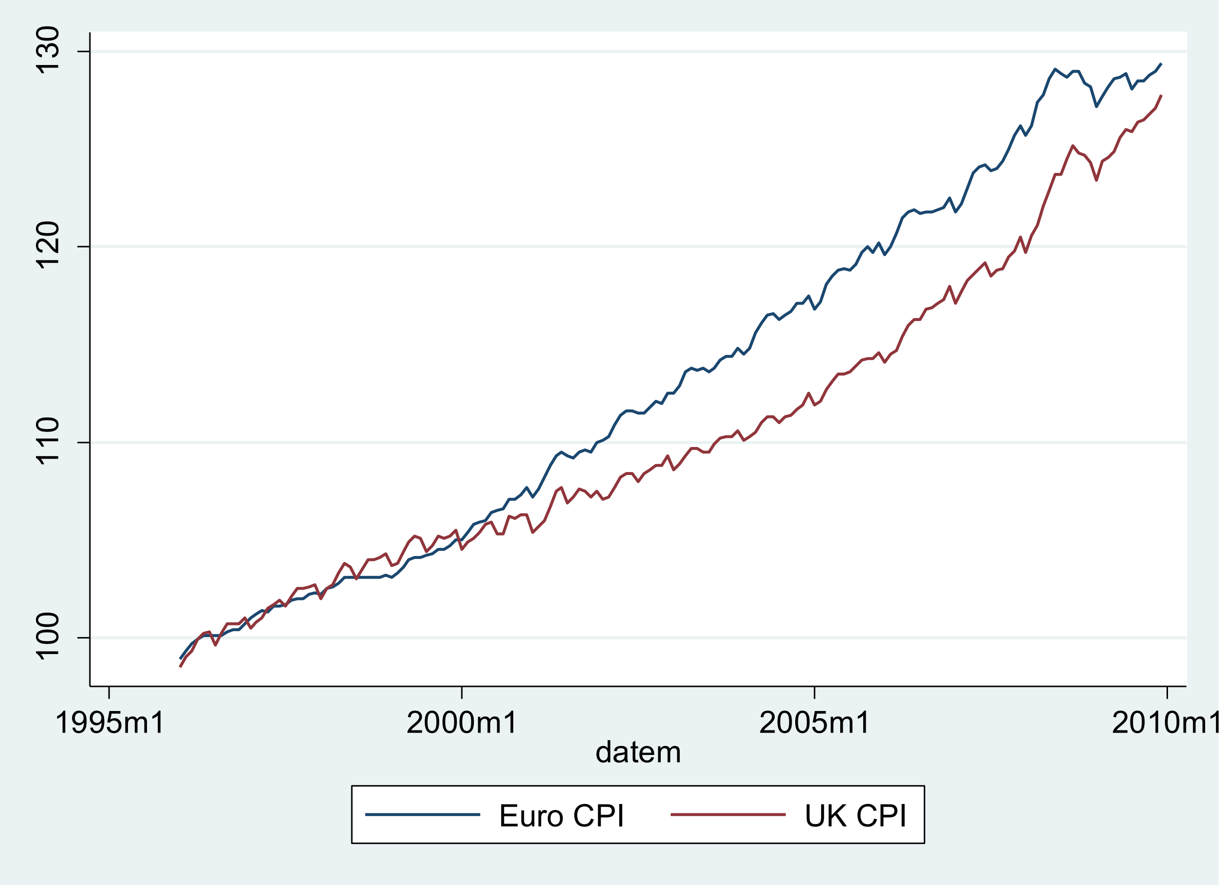
We want to look at inflation rate relationships across countries.

If we used the price index in levels, we would have more power and be able to say more…

Is that possible?



Cointegration Example



ukpi.dta

Steps:

1) Show returns are stationary (levels clearly are not!)
2) Estimate residuals from levels series
3) Test residuals for stationarity.
   - If they are, then the series are cointegrated, we know the cointegrating vector, and we are all set!
   - If they are not, they we have to use returns

Cointegration

1) Show returns are stationary (levels clearly are not!)

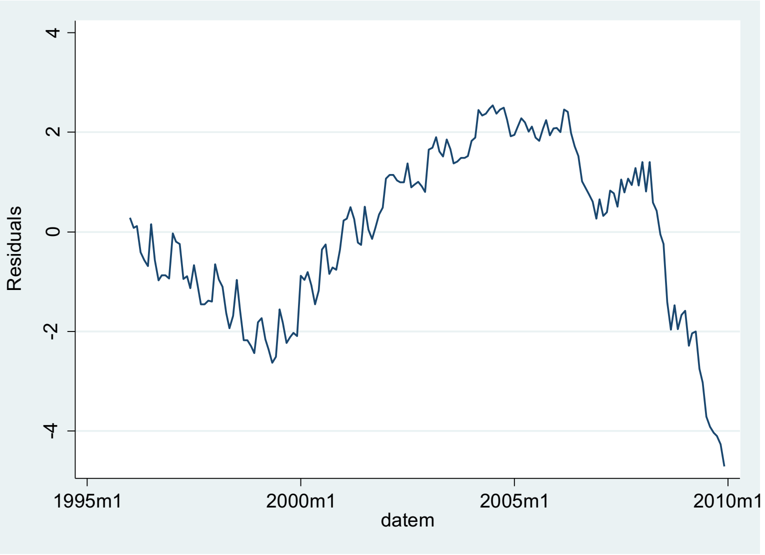
```
. gen ret_cpi_eur = d.euro/l.euro
(1 missing value generated)

. gen ret_cpi_uk = d.uk/l.uk
(1 missing value generated)

. dfuller ret_cpi_uk

Dickey-Fuller test for unit root                   Number of obs   =       166

                               ---------- Interpolated Dickey-Fuller ----------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)           -13.712            -3.488            -2.886            -2.576
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000
```

Reject the null of a unit root....

Cointegration

```
. reg euro uk

      Source |       SS           df       MS      Number of obs   =       168
-------------+----------------------------------   F(1, 166)       =   5119.24
       Model |  14544.7489         1  14544.7489   Prob > F        =    0.0000
    Residual |  471.637735       166  2.84119118   R-squared       =    0.9686
-------------+----------------------------------   Adj R-squared   =    0.9684
       Total |  15016.3866       167  89.9184827   Root MSE        =    1.6856

------------------------------------------------------------------------------
        euro |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          uk |   1.211646   .0169345    71.55   0.000     1.178211    1.245081
       _cons |  -20.73525   1.878071   -11.04   0.000    -24.44324   -17.02727
------------------------------------------------------------------------------

. predict ehat, resid

. dfuller ehat

Dickey-Fuller test for unit root                   Number of obs   =       167

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -0.085            -3.488            -2.886            -2.576
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.9509
```

1) Estimate residuals from levels series
2) Test residuals for stationarity.



Can’t say it is not non-stationary.
Can’t do cointegration (with this form)
=> Use returns model

Yule Walker Equations

If we have stationary series, we know that the moments and covariances are not functions of time (by definition).

This means that we may be able to solve for them in more straightforward ways….

Let's start with the model

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

Yule Walker Equations

Imagine that we define the means and variances/autocovariances as follows:
$\mu = E[y_t];$
$\gamma_0 = E[(y_t - \mu)(y_t - \mu)]$
$\gamma_1 = E[(y_t - \mu)(y_{t-1} - \mu)]$

And so on for any t (assuming again that such a value is defined/the moments converge).

We can write out expressions for each of these, and then solve for μ, γ_i as a function of the parameters…

Let's begin with an AR(1) example:

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

Yule Walker Equations

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

$$E[y_t] = E[c] + E[\phi y_{t-1}] + E[\varepsilon_t]$$

Plug in $\mu = y_t$ for all t

$$\mu = c + \phi\mu$$

$$\mu = \frac{c}{1-\phi}$$

Which is exactly the mean of our AR(1) series!

Yule Walker Equations

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

Subtract μ from each side:

$$y_t - \mu = c + \phi y_{t-1} + \varepsilon_t - \frac{c}{1-\phi}$$

plug in for $c = \mu(1-\phi)$:

$$y_t - \mu = \mu(1-\phi) + \phi y_{t-1} + \varepsilon_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t$$

Can use this to solve for the variance!

Yule Walker Equations

$$y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t$$

$$E[(y_t - \mu)(y_t - \mu)] = E[[\phi(y_{t-1} - \mu) + \varepsilon_t]^2]$$

Plug back in for $\gamma_0 = E[(y_t - \mu)(y_t - \mu)] = E[(y_t - \mu)^2]$ for any t

$$\gamma_0 = \phi^2\gamma_0 + E[2\phi(y_{t-1} - \mu)\varepsilon_t] + E[{\varepsilon_t}^2]$$

Yule Walker Equations

$$\gamma_0 = \phi^2\gamma_0 + E[2\phi(y_{t-1} - \mu)\varepsilon_t] + E[{\varepsilon_t}^2]$$

$$\gamma_0 = \phi^2\gamma_0 + \qquad 0 \qquad + E[{\varepsilon_t}^2]$$

$$\gamma_0 = \phi^2\gamma_0 + \sigma^2$$
$$\gamma_0 = \frac{\sigma^2}{(1 - \phi^2)}$$

Which is exactly the variance of our AR(1) series!

Yule Walker Equations

These can be much more convenient ways to describe the series and to get at the underlying relationships.

Consider an AR(2) series:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

AR(2) Models

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

Plug in for $E[y_t] = \mu$

$$\mu = c + \phi_1 \mu + \phi_2 \mu$$

$$\mu = \frac{c}{1 - \phi_1 - \phi_2}$$

AR(2) Models

Now let's solve for $\gamma_0 = E[(y_t - \mu)(y_t - \mu)]$. First, subtract off μ

$$y_t - \mu = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t - \mu$$

Plug in for $c = (1 - \phi_1 - \phi_2)\,\mu$

$$y_t - \mu = (1 - \phi_1 - \phi_2)\,\mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} - \mu + \varepsilon_t$$

Collect terms in ϕ

$$y_t - \mu = \phi_1 (y_{t-1} - \mu) + \phi_2 (y_{t-2} - \mu) + \varepsilon_t$$

Multiply by $(y_t - \mu)$, take expectations

AR(2) Models

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \varepsilon_t$$

Multiply by $(y_t - \mu)$, take expectations

$$E[(y_t - \mu)(y_t - \mu)] = E[\phi_1(y_{t-1} - \mu)(y_t - \mu)] + E[\phi_2(y_{t-2} - \mu)(y_t - \mu)] + E[\varepsilon_t(y_t - \mu)]$$

Define $\gamma_0, \gamma_1, \gamma_2$ as the first 0, 1, 2 autocovariances:

$$\gamma_0 = E[(y_t - \mu)(y_t - \mu)]$$

$$\gamma_1 = E[(y_{t-1} - \mu)(y_t - \mu)]$$

$$\gamma_2 = E[(y_{t-2} - \mu)(y_t - \mu)]$$

AR(2) Models

$$E[(y_t-\mu)(y_t-\mu)] = E[\phi_1(y_{t-1}-\mu)(y_t-\mu)] + E[\phi_2(y_{t-2}-\mu)(y_t-\mu)] + E[\varepsilon_t(y_t-\mu)]$$

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2$$

If we instead multiplied everything by $(y_{t-1}-\mu)$:

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1$$

If we instead multiplied everything by $(y_{t-2}-\mu)$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0$$

AR(2) Models

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2$$

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1$$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0$$

Three equations, three unknowns.

What do we actually have here? These are *moment restrictions* (relationships on the autocovariances). Can use them to estimate the parameters in a dataset like any other method of moments estimator.

ARMA Models

Combine AR and MA effects together:

Called an ARMA(p,q) model where p = order of AR effects, q = order of MA effects.

These can have unusual AC and PAC patterns… consider an ARMA(1,1)

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

ARMA Models

ARMA(1,1)

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

Plug in $E[y_t] = \mu$

$$\mu = \phi\mu + E\varepsilon_t + \theta E\varepsilon_{t-1} \;=>\; \mu = \phi\mu$$

$$\Rightarrow \; \mu = 0$$

ARMA Models

ARMA(1,1)

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$$

Let's see what the Yule Walker Equations look like/what we can expect our ACF and PACF to show…
First multiple both sides by $y_t - \mu = y_t$, then will repeat with $y_{t-1} - \mu$. etc …

$$E[(y_t - \mu)(y_t - \mu)] = E[\ \phi y_{t-1}y_t + \varepsilon_t y_t + \theta\varepsilon_{t-1}y_t]$$

First term is $\phi\gamma_1$ right away, second has a σ^2 in it… third one does too since y_t is a function of y_{t-1} …

ARMA Models

ARMA(1,1)

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

Let's see what the Yule Walker Equations look like/what we can expect our ACF and PACF to show...
First multiple both sides by $y_t - \mu = y_t$, then will repeat with $y_{t-1} - \mu$. etc ...

$$E[(y_t - \mu)(y_t - \mu)] = E[\ \phi y_{t-1} y_t + \varepsilon_t y_t + \theta \varepsilon_{t-1} y_t]$$

Substitute in for last y_t to have terms in t-1:

$$\gamma_0 = \phi \gamma_1 + \sigma_\varepsilon^2 + E[\theta \varepsilon_{t-1}\ (\phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1})]$$

ARMA Models

ARMA(1,1)

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

$$\gamma_0 = \phi \gamma_1 + \sigma_\varepsilon^2 + E[\theta \varepsilon_{t-1} (\phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1})]$$

Substitute out y_{t-1} again to get all of the t-1 terms for epsilon:

$$= \phi \gamma_1 + \sigma_\varepsilon^2 + E[\theta \varepsilon_{t-1} (\phi^2 y_{t-2} + \phi \varepsilon_{t-1} + \theta \phi \varepsilon_{t-2} + \varepsilon_t + \theta \varepsilon_{t-1})]$$

$$= \phi \gamma_1 + \sigma_\varepsilon^2 + \theta \phi \sigma_\varepsilon^2 + \theta^2 \sigma_\varepsilon^2$$

$$\boldsymbol{\gamma_0 = \phi \gamma_1 + \sigma_\varepsilon^2 (1 + \theta \phi + \theta^2)}$$

ARMA Models

ARMA(1,1)
Now multiply both sides by $y_{t-1} - \mu = y_{t-1}$ to get an expression for γ_1

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

$$\gamma_1 = E[(y_t - \mu)(y_{t-1} - \mu)] = E[\, \phi y_{t-1} y_{t-1} + \varepsilon_t y_{t-1} + \theta \varepsilon_{t-1} y_{t-1}]$$

$$\gamma_1 = \phi \gamma_0 + E[\varepsilon_t y_{t-1}] + E[\, \theta \varepsilon_{t-1}[\phi y_{t-2} + \varepsilon_{t-1} + \theta \varepsilon_{t-2}]]$$

$$\boldsymbol{\gamma_1 = \phi \gamma_0 + \theta \sigma_\varepsilon^2}$$

ARMA Models

ARMA(1,1)
And again for y_{t-2}

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

$$\gamma_2 = E[(y_t - \mu)(y_{t-2} - \mu)] = E[\phi y_{t-1}\, y_{t-2} + \varepsilon_t y_{t-2} + \theta \varepsilon_{t-1} y_{t-2}]$$

$$\boldsymbol{\gamma_2 = \phi \gamma_1}$$

This is an AR-style exponential decay. Gets smaller by a multiple of ϕ each period

ARMA Models

$$\gamma_0 = \phi\gamma_1 + \sigma_\varepsilon^2(1 + \theta\phi + \theta^2)$$
$$\gamma_1 = \phi\gamma_0 + \theta\sigma_\varepsilon^2$$
$$\gamma_2 = \phi\gamma_1$$

ARMA(p,q) models => initial q periods have complex autocovariances, simplify to AR style after that

ARMA Models

MATLAB example

```
mu = 0;
phi = [0.8;-0.5]; %AR COEFFICIENTS
lambda = -0.3; %MA COEFFICIENTS

T = [1:1:1000]';
%epsilon is mean zero, variance one
eps = mvnrnd(0,1,1000);

%to store the simulation

y_ar2 = zeros(1000,1);
y_arma11 = zeros(1000,1);
y_ar1 = zeros(1000,1);

%run the simulation period by period
for i = 3:1000
    y_ar1(i) = phi(1)*y_ar1(i-1) + eps(i);
    y_ar2(i) = phi(1)*y_ar2(i-1)+ phi(2)*y_ar2(i-2) +eps(i);
    y_arma11(i) = phi(1)*y_arma11(i-1)+ eps(i)+lambda*eps(i-1);
end
```

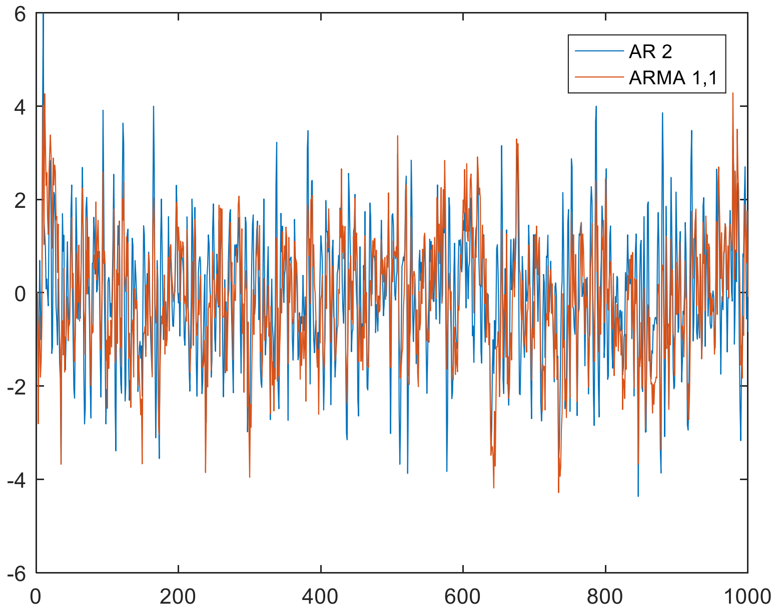
ARMA Models

MATLAB example

```
%run the simulation period by period
for i = 3:1000
    y_ar1(i) = phi(1)*y_ar1(i-1) + eps(i);
    y_ar2(i) = phi(1)*y_ar2(i-1)+ phi(2)*y_ar2(i-2) +eps(i);
    y_arma11(i) = phi(1)*y_arma11(i-1)+ eps(i)+lambda*eps(i-1);
end


%plot the series
plot(T,[y_ar2,y_arma11])
legend('AR 2','ARMA 1,1')

%AC and PAC plots…. THINK ABOUT WHAT THESE SHOULD LOOK LIKE FOR EACH
autocorr(y_ar1)
parcorr(y_ar1)

autocorr(y_ar2)
parcorr(y_ar2)

autocorr(y_arma11)
parcorr(y_arma11)
```

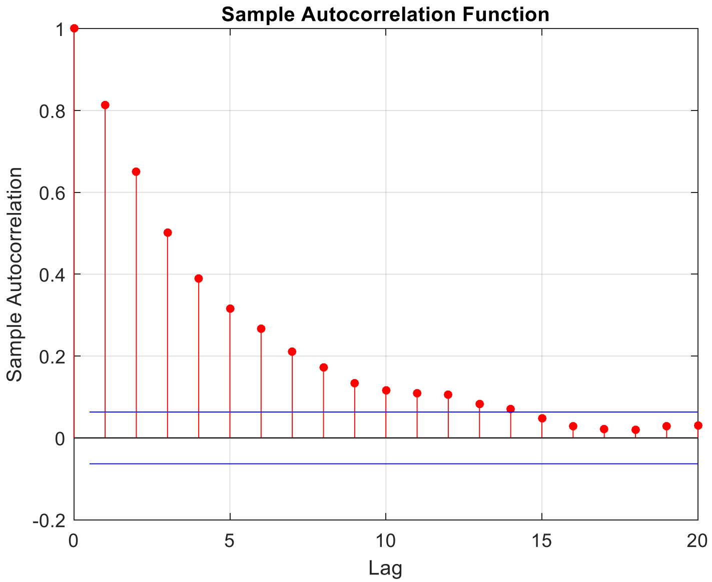
ARMA Models

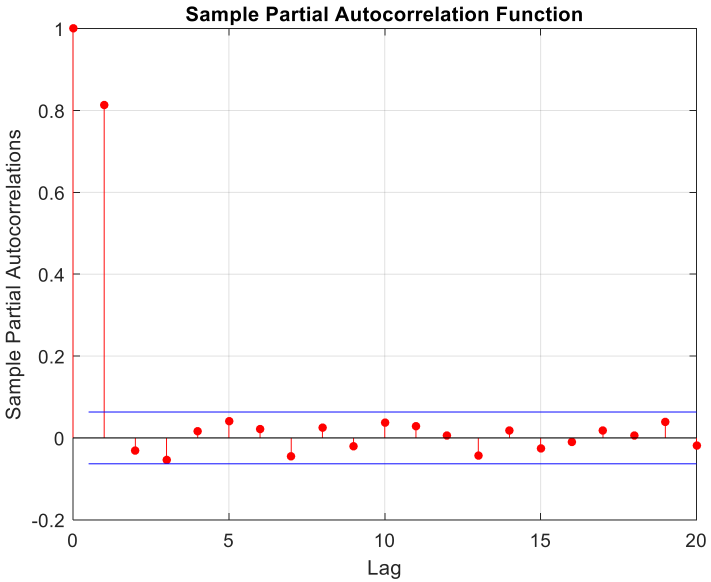
```
plot(T,[y_ar2,y_arma11])
legend('AR 2','ARMA 1,1’)
```



ARMA Models

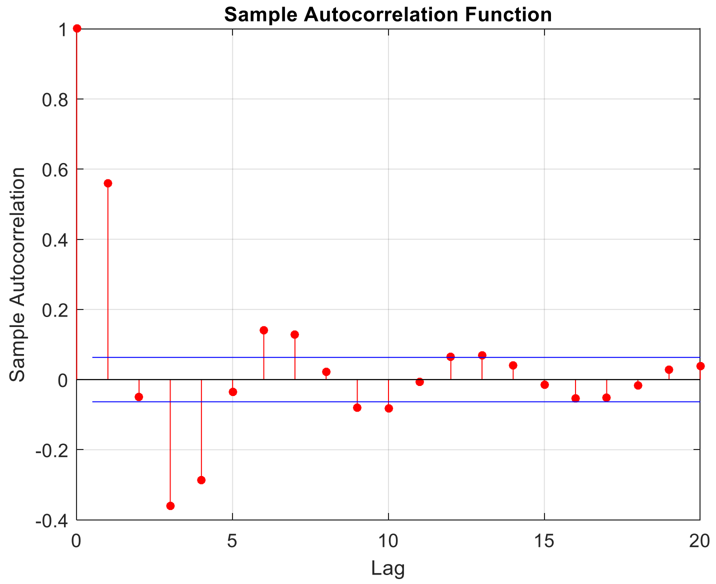
```
autocorr(y_ar1)
parcorr(y_ar1)
```

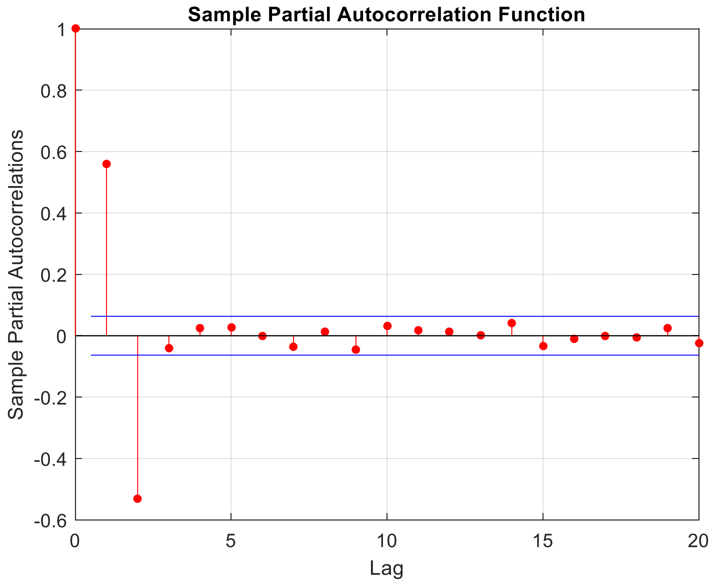




ARMA Models

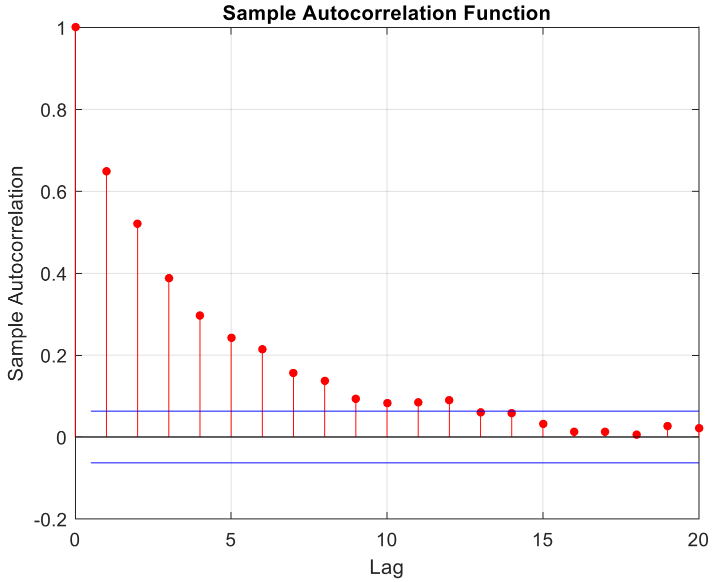
```
autocorr(y_ar2)
parcorr(y_ar2)
```

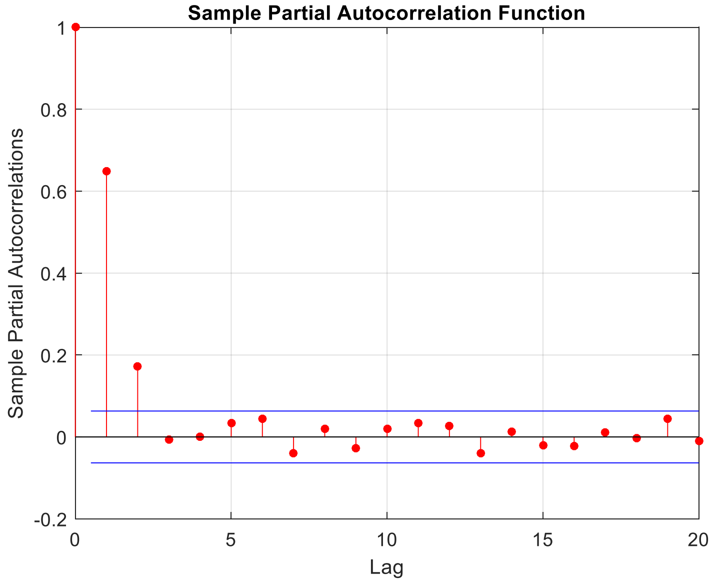




ARMA Models

```
autocorr(y_arma11)
parcorr(y_arma11)
```





ARMA Models

Not always obvious what you are working with from the plots.

This one could be AR(2) or ARMA(1,1)… of course, we know what the right answer is, but you will never be so lucky.

Keys:

- Do we have a reason to expect something here? Seasonality, business cycles?
- Simpler is usually better. Long dependencies are hard to justify
- Implications matter. If one gives you much tighter forecast bands and looks substantively similar, that's a good argument