Introduction to Time Series Models

TS Data Analysis

- Graphical (vs time)
  - level or change
  - single or multiple
- Numerical
  - Decomposition
  - Generating Process

Covered Topics

- Decomposition Approach (only introduction)
- Generating Process Approach (main topic)
- Difference Equation

Definition

Define Y as a time series

Y_t is the observed value of Y at time t. Note that Y_t is a random variable

t is used as the observation index (in place of i) which indicates the chronological order.

Components of TS (1)

- Time trend as function of t (Tr_t)
- Seasonal (Sn_t)
- Cyclical (Cl_t)
- Irregular or random (Ir_t)

Not to be discussed in details

Components of TS (2)

Some patterns of time trend (Tr_t)

- linear
- polynomial
- lin-log, log-lin, log-log
- reciprocal
- logistic

Components of TS (3)

Modelling the TS components

- Additive

$$Y_t = Tr_t + Sn_t + Cl_t + Ir_t$$

- Multiplicative

$$Y_t = Tr_t * Sn_t * Cl_t * Ir_t$$

Decomposition Method (1)

Steps

- ignore the Cyclical component
- identify and separate the trend (de-trend)
- identify and separate seasonal component from the de-trended data

Decomposition Method (2)

To separate components, use averaging and/or regression and/or smoothing technique

To forecast Y, forecast each component and, then, add or multiply them, depending on the model selected.

Smoothing Non-seasonal

- Single-exponential Smoothing
- Double-exponential Smoothing
- Triple-exponential Smoothing
- Holt-Winters's exponential Smoothing (2 smoothing constants)

Smoothing Seasonal

- Holt-Winters's exponential Smoothing (3 smoothing constants)

See EViews for more explanation.

Note that Eviews does not support triple-smoothing.

TS Generating Processes

- Linear in t, Y_1, Y_2,…, Y_t,… and random components
- Non-linear in t, Y_1, Y_2,…, Y_t, … and random components

Linear Generating Processes

- White noise
- Auto-regressive (AR)
- Moving average (MA)
- Mixed AR and MA (ARMA)
- Differencing (ARIMA)
- ARIMAX (ARIMA with X)

White Noise Processes

Note that *Y* is a white noise process if

$$E(Y_t) = 0 \quad \text{for } \forall t$$

$$V(Y_t) = \sigma^2 \quad \text{for } \forall t$$

$$Cov(Y_t, Y_s) = 0 \quad \text{for } \forall t \neq s$$

That is, $Y_t = \varepsilon_t$.

AR Processes

Basic Form of AR(p) or AR process of order p

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

where ε_t is a white noise.

That is, Y_t. is a linear function of its own lagged values plus a random component.

MA Processes

Basic Form of MA(q) or MA process of order q

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

where ε's are white noises

That is, Y_t. is a moving average of the current and lagged values of white noses.

ARMA Processes

Basic Form of ARMA(p,q)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$$
$$+ \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$
$$+ \cdots + \theta_q \varepsilon_{t-q}$$

Note that

AR(p)=ARMA(p,0)

MA(q) = ARMA(0,q)

AR(0)=MA(0)=white noise

1st & 2nd Differences

1st difference

$$\Delta^1 Y_t \equiv Y_t - Y_{t-1}$$

2nd difference

$$\Delta^2 Y_t \equiv \Delta Y_t - \Delta Y_{t-1}$$
$$\equiv Y_t - 2Y_{t-1} + Y_{t-2}$$

Higher Differences

3rd difference

$$\Delta^3 Y_t \equiv \Delta(\Delta^2 Y_t) - \Delta(\Delta^2 Y_{t-1})$$

$$\equiv \left(Y_t - 2Y_{t-1} + Y_{t-2}\right) - \left(Y_{t-1} - 2Y_{t-2} + Y_{t-3}\right)$$

$$\equiv Y_t - 3Y_{t-1} + 3Y_{t-2} - Y_{t-3}$$

dth difference

$$\Delta^d Y_t \equiv \Delta(\Delta^{d-1} Y_t)$$

$$\equiv \Delta^{d-1} Y_t - \Delta^{d-1} Y_{t-1}$$

Note that $\Delta^0 Y_t \equiv Y_t$

Seasonal Differences

If there are S seasons in a year,
1st seasonal difference

$$\Delta_S Y_t = Y_t - Y_{t-S}$$

Note that $\Delta_4 Y_t \neq \Delta^4 Y_t$

dth seasonal difference

$$\Delta_S^d Y_t = \Delta_S^d (\Delta_S^{d-1} Y_t)$$

$$= \Delta_S^{d-1} Y_t - \Delta_S^{d-1} Y_{t-S}$$

ARIMA Processes

Basic Form of ARIMA(p,d,q)

$$\Delta^d Y_t = \phi_1 \Delta^d Y_{t-1} + \phi_2 \Delta^d Y_{t-2} + \cdots + \phi_p \Delta^d Y_{t-p}$$
$$+ \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

I stands for Integrated. d is called the integrated order.

Note that ARIMA(p,0,q)=ARMA(p,q)

ARIMAX Processes

$$\Delta^d Y_t = \beta_1 X_{1t} + \cdots + \beta_K X_{Kt}$$
$$+ \phi_1 \Delta^d Y_{t-1} + \phi_2 \Delta^d Y_{t-2}$$
$$+ \cdots + \phi_p \Delta^d Y_{t-p}$$
$$+ \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

Note that some of the X's could be a constant or time trend or dummy variables.

Key Issues

- Parameter Estimation
  - Is LS valid?
  - Any additional assumption?
- Their Standard errors
- Predict Y_{t+i}

Difference Equation

- Discrete analogy to differential equations
- Describe how a quantitative variable (Y) is related to its past values and exogenous series (x)
- Basis for a time series model

General Form

$$F(\Delta^p Y_t, \Delta^{p-1} Y_{t-1}, \ldots, \Delta^0 Y_{t-p}, x_t) = 0$$

or $F(\Delta^p Y_t, \Delta^{p-1} Y_t, \ldots, \Delta^0 Y_t, x_t) = 0$

Linear Form

$$\Delta^p Y_t + \beta_1 \Delta^{p-1} Y_{t-1} + \beta_2 \Delta^{p-2} Y_{t-2} + \cdots + \beta_p \Delta^0 Y_{t-p} = x_t$$

More Preferable Form

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + x_t$$

where $\phi_1 = p - \beta_1$

$$\phi_2 = -\frac{1}{2} p(p-1) + (p-1)\beta_1 - \beta_2$$

$$\vdots$$

Lag Operator

$$LY_t \equiv Y_{t-1}$$

$$L^2 Y_t \equiv L(LY_t) \equiv L(Y_{t-1}) \equiv Y_{t-2}$$

$$\vdots$$

$$L^p Y_t \equiv L(L^{p-1} Y_t) \equiv L(Y_{t-(p-1)}) \equiv Y_{t-p}$$

Basic algebra tool for TS process analyses

Alternative Form

Using Lag Operator

$$(1-\phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) Y_t = x_t$$

$$(1-\lambda_1 L)(1-\lambda_2 L)\ldots(1-\lambda_p L) Y_t = x_t$$

where λ's are p characteristic roots of

$$\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \cdots - \phi_p \lambda^0 = 0$$

$$(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_p) = 0$$

Solve Diff. Eqn

Determine the values of

$$Y_t, Y_{t-1}, Y_{t-2}, \ldots$$

in terms of $\phi_1, \cdots, \phi_p, t, u_t, u_{t-1}, \ldots$

that satisfies the difference equation

Note that u could be constant, time trend or another time series

Solution (1)

Homogenous Solution u_t=0

$$Y_t^h = \sum_{j=1}^{p} A_j \lambda_j^t$$

Particular Solution Y_t=Y_{t-1}

$$Y_t^P = \frac{1}{1-\phi_1-\phi_2-\cdots-\phi_p} x_t$$

Solution (2)

General Solution

$$Y_t = Y_t^h + Y_t^p$$
$$= \sum_{j=1}^{p} A_j \lambda_j^t + \frac{1}{1-\phi_1-\phi_2-\cdots-\phi_p} x_t$$

Solution (3)

Homo. solution is convergent or stable if (exact conditions)

$$\left|\lambda_j\right| < 1 \text{ for } \forall j$$

=> Y_t^h converges to zero

Non-convergent if

$$\left|\lambda_j\right| \geq 1 \text{ for some } j$$

Solution (4)

Necessary conditions

$$\sum_{j=1}^{p} \phi_j < 1$$

Sufficient Conditions

$$\sum_{j=1}^{p} \left| \phi_j \right| < 1$$

Schur Theorem

Schur Theorem (1)

$$\Delta_1 = \begin{vmatrix} 1 & -\phi_p \\ -\phi_p & 1 \end{vmatrix} > 0$$

$$\Delta_2 = \begin{vmatrix} 1 & 0 & -\phi_p & -\phi_{p-1} \\ -\phi_1 & 1 & 0 & -\phi_p \\ -\phi_p & 0 & 1 & -\phi_1 \\ -\phi_{p-1} & -\phi_p & 0 & 1 \end{vmatrix} > 0$$

Schur Theorem (2)

$$\Delta_3 = \begin{vmatrix} 1 & 0 & 0 & -\phi_p & -\phi_{p-1} & -\phi_{p-2} \\ -\phi_1 & 1 & 0 & 0 & -\phi_p & -\phi_{p-1} \\ -\phi_2 & -\phi_1 & 1 & 0 & 0 & -\phi_p \\ -\phi_p & 0 & 0 & 1 & -\phi_1 & \\ -\phi_{p-1} & -\phi_p & 0 & 0 & 1 & -\phi_1 \\ -\phi_{p-2} & -\phi_{p-1} & -\phi_p & 0 & 0 & 1 \end{vmatrix} > 0$$

Schur Theorem (3)

$$\Delta_p = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & -\phi_p & -\phi_{p-1} & \cdots & -\phi_2 & -\phi_1 \\ & 1 & \ddots & \ddots & 0 & 0 & -\phi_p & \ddots & \ddots & -\phi_2 \\ & & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ & & & 1 & 0 & 0 & \ddots & \ddots & -\phi_p & -\phi_{p-1} \\ & & & & 1 & 0 & 0 & \cdots & 0 & -\phi_p \\ & & & & & 1 & & & & \\ & & & & & 0 & 1 & & & \\ & & & & & \vdots & \ddots & \ddots & & \\ & & & & & 0 & \ddots & \ddots & 1 & \\ & & & & & 0 & 0 & \cdots & 0 & 1 \end{vmatrix} > 0$$

Auto-correlation of Error Terms

Pongsa Pornchaiwiseskul
Faculty of Economics
Chulalongkorn University

Covered Topics

- GLS
- Assumptions
- Detection, e.g.,
  - Durbin-Watson d-test
  - Breusch-Godfrey test
- Estimation Methods, e.g.,
  - Cochrane-Orcutt

General Auto-correlation (1)

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\nu}$$

$$\mathbf{E}(\boldsymbol{\nu}) = \mathbf{0}$$

$$\mathbf{V}(\boldsymbol{\nu}) = \sigma^2\boldsymbol{\Sigma}$$

where

0 is a nx1 column vector of zeroes

Σ is an nxn <u>positive-definite symmetric</u> matrix.

General Auto-correlation (2)

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdot & \sigma_{1n} \\ \sigma_{12} & \sigma_{22} & \cdot & \sigma_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{1n} & \sigma_{2n} & \cdot & \sigma_{nn} \end{bmatrix}$$

where $\sigma^2\sigma_{ij} = \mathrm{V}(\nu_i, \nu_j)$ for $i, j = 1, \ldots, n$

General Auto-correlation (3)

Sources of Auto-correlation

- Inertia (nature)
- Spill-over effect over geographical region, e.g., contagion, migration (nature)
- Spec. errors, e.g.,
  - Exclusion of auto-correlated independent variables
  - Incorrect functional form
  - Lagged terms

General Auto-correlation (4)

Effect of Ignoring Pure Auto-correlation

- OLS is unbiased but not the best
- Need new estimate for Σ, e.g.,
  - Newey-West formula

Generalized LS (1)

$$\mathbf{\Omega Y} = \mathbf{\Omega X\beta} + \mathbf{\Omega \nu}$$

where Ω is an nxn symmetric matrix such that

$$\mathbf{\Omega\Omega} = \mathbf{\Sigma}^{-1}$$

Note that E($\Omega\nu$)=**0** and V($\Omega\nu$)=$\sigma^2\mathbf{I}_n$

Generalized LS (2)

If Σ is known, apply OLS to BLU estimate β,σ^2

$$\hat{\boldsymbol{\beta}} = \left[(\mathbf{\Omega X})^T(\mathbf{\Omega X})\right]^{-1}(\mathbf{\Omega X})^T(\mathbf{\Omega Y})$$

$$= \left[\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X}\right]^{-1}\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{Y}$$

$$\hat{\sigma}^2 = \tfrac{1}{n-K}\left(\mathbf{\Omega}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})\right)^T\left(\mathbf{\Omega}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})\right)$$

$$= \tfrac{1}{n-K}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T\mathbf{\Sigma}^{-1}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})$$

What if Σ is unknown? Need more assumptions?

Assumptions

$$Y_i = X_{1i}\beta_1 + X_{2i}\beta_2 + \ldots + X_{Ki}\beta_K + \nu_i$$

$$i = 1, \ldots, n$$

In addition to CLRM assumptions

- ARMA error term (ν)
- weakly stationary error term. Otherwise, estimation will be invalid. Why?

AR(1) error term --(1)

$$Y_t = X_{1t}\beta_1 + X_{2t}\beta_2 + \ldots + X_{Kt}\beta_K + \rho\nu_{t-1} + \varepsilon_t$$

$$-1 < \rho < 1, \quad t = 1, \ldots, n$$

ε_t ~ White Noise

$\mathrm{H}_0 : \rho = 0$

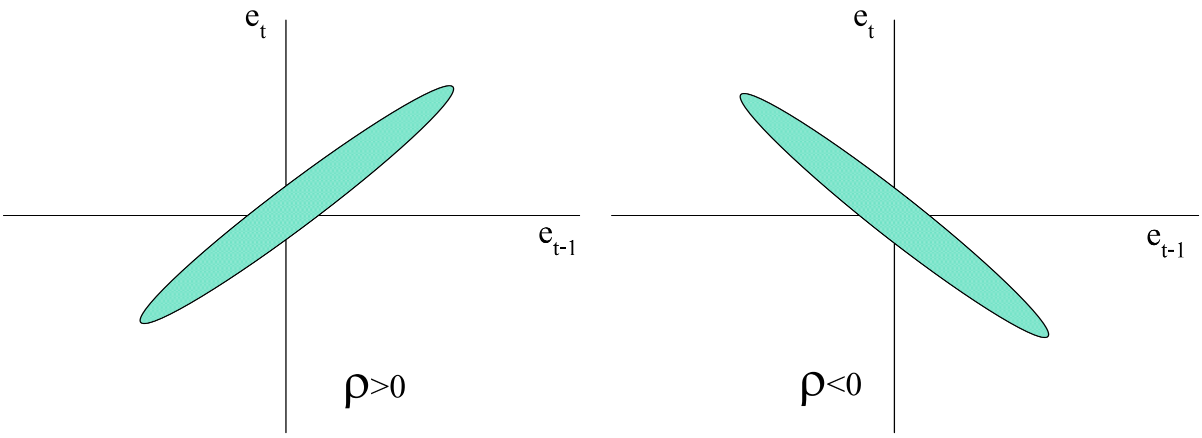
$\mathrm{H}_1 : \rho \neq 0$

Accept ==> No auto-correlation

Reject ==> AR(1)

AR(1) error term --(2)

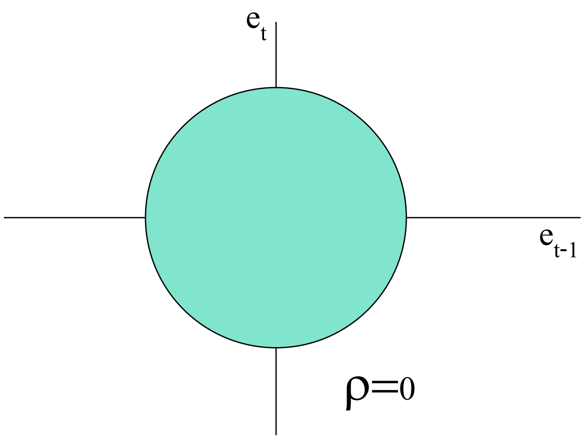
Graphical Test



where ***e*** is the OLS residual when auto-correlation has been ignored

AR(1) error term --(3)

Graphical Test (cont'd)



Note that plot of e_t vs e_{t-1} cannot reveal higher AR in error terms

AR(p) error term --(1)

$$Y_t = X_{1t}\beta_1 + X_{2t}\beta_2 + \ldots + X_{Kt}\beta_K + \rho\nu_t$$

$$\nu_t = \rho_1\nu_{t-1} + \rho_2\nu_{t-2} + \ldots + \rho_p\nu_{t-p} + \varepsilon_t$$

~ stationary

ε_t ~ White Noise

$\mathrm{H}_0 : \rho_1 = \rho_2 = \ldots = \rho_p = 0$

$\mathrm{H}_1 : \rho_1 \neq \rho_2 \neq \ldots \neq \rho_p \neq 0$

Accept ==> No auto-correlation

Reject ==> AR(p) or lower

Auto-correlation Tests

Statistical Tests

- Durbin-Watson's d-test. Good only for AR(1).
- DW's h-test (obsolete)
- Breusch-Godfrey's General Auto-correlation or Serial Correlation test
- Runs test (non-parametric). Not require normal error term nor ARMA form. Good for small sample.

Durbin-Watson's d-test (1)

$$DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

Note that t=2

DW statistic follows a special distribution. Need a <u>special table</u> to perform AR(1) test.

See Table D.5A-B (Gujarati)

Durbin-Watson's d-test (2)

$$DW = \frac{\sum_{t=2}^{n}(e_t^2 - 2e_te_{t-1} + e_{t-1}^2)}{\sum_{t=1}^{n} e_t^2}$$

$$= \frac{\sum_{t=2}^{n} e_t^2}{\sum_{t=1}^{n} e_t^2} - 2\frac{\sum_{t=2}^{n} e_te_{t-1}}{\sum_{t=1}^{n} e_t^2} + \frac{\sum_{t=2}^{n} e_{t-1}^2}{\sum_{t=1}^{n} e_t^2}$$

$$\approx 1 - 2\hat{\rho} + 1 = 2(1-\hat{\rho})$$

==> $0 \le DW \le 4$

Durbin-Watson's d-test (3)

Given significant level(α) and sample size (n), read d_L and d_U from DW tables.

Criterion:

$$0 \le DW \le d_L \;\; => \;\; \rho > 0$$

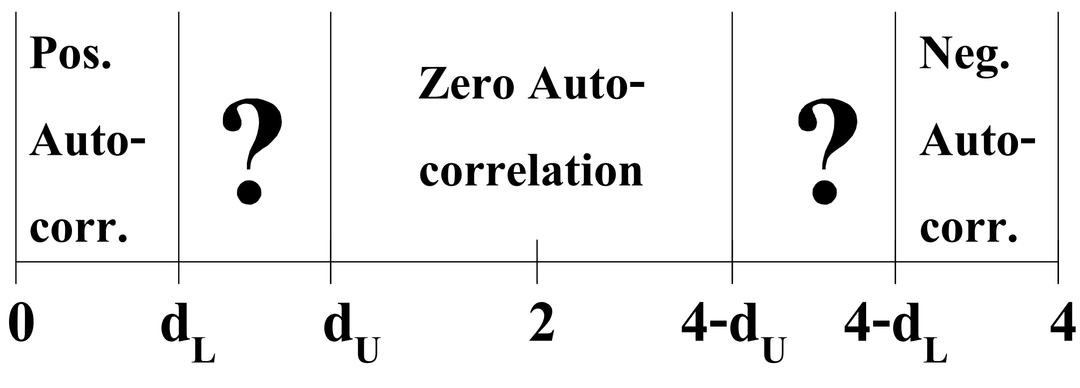
$$d_L < DW < d_U \;\; => \;\; \rho \ge 0 \quad \text{indicisive}$$

$$d_U \le DW \le 4 - d_U \;\; => \;\; \rho = 0$$

$$4 - d_U < DW < 4 - d_L \;\; => \;\; \rho \le 0 \quad \text{indicisive}$$

$$4 - d_L \le DW \le 4 \;\; => \;\; \rho < 0$$

Durbin-Watson's d-test (4)



Note that the conclusion is not simply "Accept" or "Reject".

Durbin-Watson's d-test (4)

Notes

- DW is invalid if there are lagged dependent variables as explanatory variables (AR Model). Their existence is equivalent to higher AR of error terms
- Good only for AR(1). DW statistic is usually an item in the OLS report.
- DW can be used to roughly estimate ρ but no SE given.

Breusch-Godfrey's test (1)

Step1 Run OLS by ignoring auto-correlation

==> residual $\hat{v}_t$

Step 2 run OLS for

$$\hat{v}_t = X_{1t}\beta_1 + X_{2t}\beta_2 + \ldots + X_{Kt}\beta_K + \rho_1\hat{v}_{t-1} + \rho_2\hat{v}_{t-2} + \ldots + \rho_p\hat{v}_{t-p} + \varepsilon_t'$$

Breusch-Godfrey's test (2)

Step 3 Do F-Test or χ^2 test for

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_p = 0$$

$$H_1 : \rho_1 \neq \rho_2 \neq \ldots \neq \rho_p \neq 0$$

Accept H_0 => no AR(p) auto-corr.

Reject ==> AR(p) or lower

Note that BG test also gives estimate for ρ's and their SE's but still not the best estimates.

Breusch-Godfrey's test (3)

$$F_{cal} = \frac{(RSS_R - RSS_U) \div p}{RSS_U \div (n - K - p)}$$

$$= \frac{(TSS_U - RSS_U) \div p}{RSS_U \div (n - K - p)}$$

$$= \frac{n - K - p}{p} R^2 \sim F(p, n - K - p)$$

$$\chi^2{}_{cal} = (n - p) R^2 \sim \chi^2(p)$$

Cochrane-Orcutt Method (1)

For AR(p) error term

<u>Step1</u> Run OLS by ignoring auto-correlation

==> residual $\hat{v}_t$

<u>Step 2</u> run OLS for

$$\hat{v}_t = \rho_1 \hat{v}_{t-1} + \rho_2 \hat{v}_{t-2} + \ldots + \rho_p \hat{v}_{t-p} + \varepsilon'_t$$

$$==> \hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_p$$

<u>Step 3</u> Transform ***Y*** and ***X's***

Cochrane-Orcutt Method (2)

<u>Step 3</u> Transform ***Y*** and ***X's***

$$Y_t^* = Y_t - \hat{\rho}_1 Y_{t-1} - \hat{\rho}_2 Y_{t-2} - \ldots - \hat{\rho}_p Y_{t-p}$$

$$X_{kt}^* = X_{kt} - \hat{\rho}_1 X_{k,t-1} - \hat{\rho}_2 X_{k,t-2} - \ldots - \hat{\rho}_p X_{k,t-p}$$

$$k = 1, .., K$$

<u>Step 4</u> Run OLS for

$$Y_t^* = X_{1t}^* \beta_1 + X_{2t}^* \beta_2 + \ldots + X_{Kt}^* \beta_K + v_t^*$$

$$v_t^* \approx v_t - \rho_1 v_{t-1} - \rho_2 v_{t-2} - \ldots - \rho_p v_{t-p}$$

$$= \varepsilon_t$$

Cochrane-Orcutt Method (3)

$$V(v_t^*) \approx V(\varepsilon_t) = \sigma^2$$

==>OLS is almost BLUE

Step 5 Re-calculate $\hat{v}_t$ using new $\hat{\boldsymbol{\beta}}$

$$\hat{v}_t = Y_t - X_{1t}\hat{\beta}_1 - X_{2t}\hat{\beta}_2 - \ldots - X_{Kt}\hat{\beta}_K$$

If solution does not significantly change, stop. Otherwise, go back to Step 2

MA Part Estimation (1)

Note that

$$v_t^* = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

$$Cov(v_t^*, v_t^*) = \sigma^2(\theta_1^2 + \theta_2^2 + \ldots + \theta_q^2)$$

$$Cov(v_t^*, v_{t-1}^*) = \sigma^2(\theta_1\theta_2 + \theta_2\theta_3 + \ldots + \theta_{q-1}\theta_q)$$

$$\vdots$$

$$Cov(v_t^*, v_{t-q}^*) = \sigma^2(\theta_1\theta_q)$$

MA Part Estimation (2)

$$\theta_1\theta_2 + \theta_2\theta_3 + \cdots + \theta_{q-1}\theta_q = \hat{\rho}_1(\theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

$$\theta_1\theta_3 + \theta_2\theta_4 + \cdots + \theta_{q-2}\theta_q = \hat{\rho}_2(\theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

$$\vdots$$

$$\theta_1\theta_{q-1} + \theta_2\theta_q = \hat{\rho}_{q-2}(\theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

$$\theta_1\theta_q = \hat{\rho}_{q-1}(\theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

q unknown q equations ==> solve for $\hat{\boldsymbol{\theta}}$

Add as step 4.1 in Cochrane-Orcutt