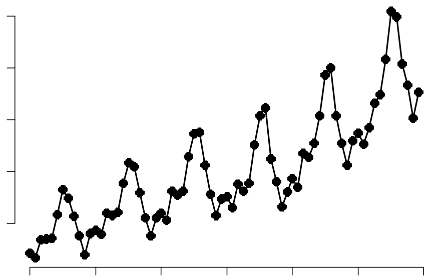


8. Multivariate Time Series

Univariate Time Series

- SARIMA works great for *causal* TS (i.e. $\rho(h) \sim \alpha^h$ as $h \to \infty$)
  - Addresses most univariate TS problems
- Cases where SARIMA fails & alternatives
  - Discrete-valued TS ⇨ Markov Chains
  - TS w/ *long memory*, i.e. $\rho(h) \sim 1/h$, ⇨ fractional integration
  - TS w/ *stochastic*, i.e. non-constant, variance ⇨(G)ARCH models

Multivariate Time Series

- When analyzing *multivariate* TS, there are many more interesting questions to ask:
  - Estimation & Model selection
  - Prediction (a.k.a. Forecasting)

same as univariate

  - Are different TS's related?
  - Does one TS *lead* the other(s)
  - How do changes in one affect other(s)

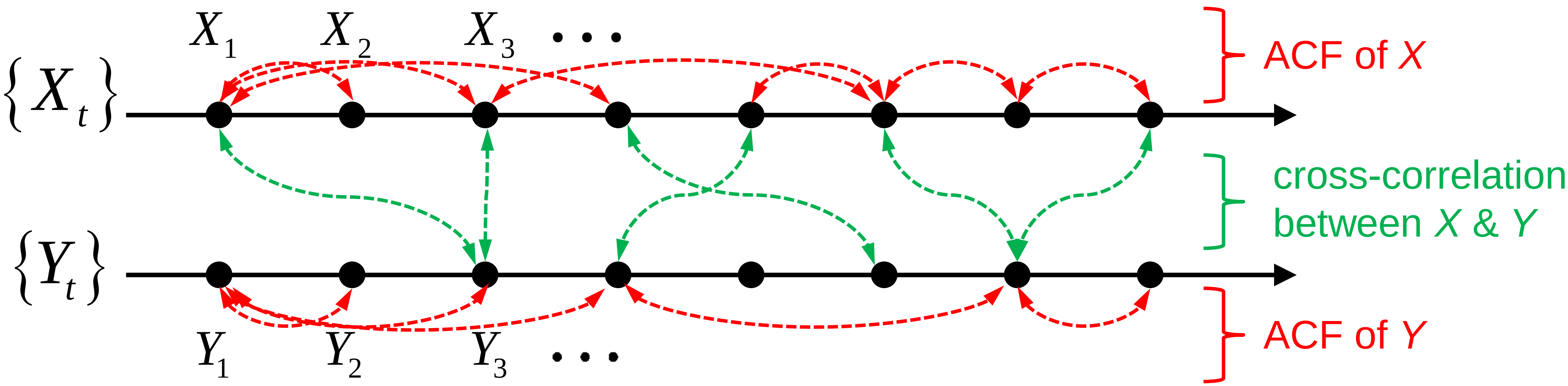
multivariate only

Multivariate Time Series

- Consider two TS $\{X_t\}, \{Y_t\}$. For *prediction* we can fit separate SARIMA models, so why bother with bivariate analysis?
  - If TS are independent / uncorrelated, then we can safely look at them separately
  - If TS are dependent, then predictions are better looked at jointly
- Cross-covariance looks at linear dependence of pairs of TS

Cross-Covariance

- Consider bivariate TS with components $\{X_t, Y_t\}$



  - To study dependence *within* X_t or Y_t , look at their respective autocovariances / ACFs
  - To study dependence *between* X_t *and* Y_t, look at *cross-covariance / cross-correlation* (CCF) function

Cross-Covariance / Cross-Correlation

- *Cross-covariance function* between X_t & Y_t

$$\gamma_{XY}(s,t) = Cov[X_s, Y_t] = E[(X_s - \mu_{X_s})(Y_t - \mu_{Y_t})]$$

- *Cross-correlation function* (*CCF*) b/t X_t & Y_t

$$\rho_{XY}(s,t) = \frac{\gamma_{XY}(s,t)}{\sqrt{\gamma_X(s,s)\gamma_Y(t,t)}}, \quad \left(\begin{array}{c}\text{where } \gamma_X(s,t) \text{ is} \\ \text{auto-cov. of } X_t\end{array}\right)$$

  - Note: $\gamma_{XY}(s,t) = \gamma_{YX}(t,s)$ & similarly for $\rho_{XY}(s,t)$
  - For trivariate TS $\{X_t, Y_t, Z_t\}$, would look at *all pair-wise* cross-covariances (γ_{XY}, γ_{XZ}, γ_{YZ}), & so on…

Joint Stationarity

- Two TS $\{X_t, Y_t\}$ are called *jointly stationary* if:
  1. X_t and Y_t are each stationary
  2. The cross-covariance is a function of $h = s - t$

$$\gamma_{XY}(h) = Cov[X_{t+h}, Y_t] = E[(X_{t+h} - \mu_X)(Y_t - \mu_Y)]$$

for $h = 0, \pm 1, \pm 2, \ldots$

- CCF of jointly stationary TS becomes

$$\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}} = \frac{\gamma_{XY}(h)}{\sqrt{\sigma_X^2 \sigma_Y^2}}$$

Auto- vs Cross-Covariance

- For stationary $\{X_t\}$, autocov. is *symmetric*:

$$\gamma(h) = Cov[X_{t+h}, X_t] = Cov[X_t, X_{t+h}] = \gamma(-h)$$

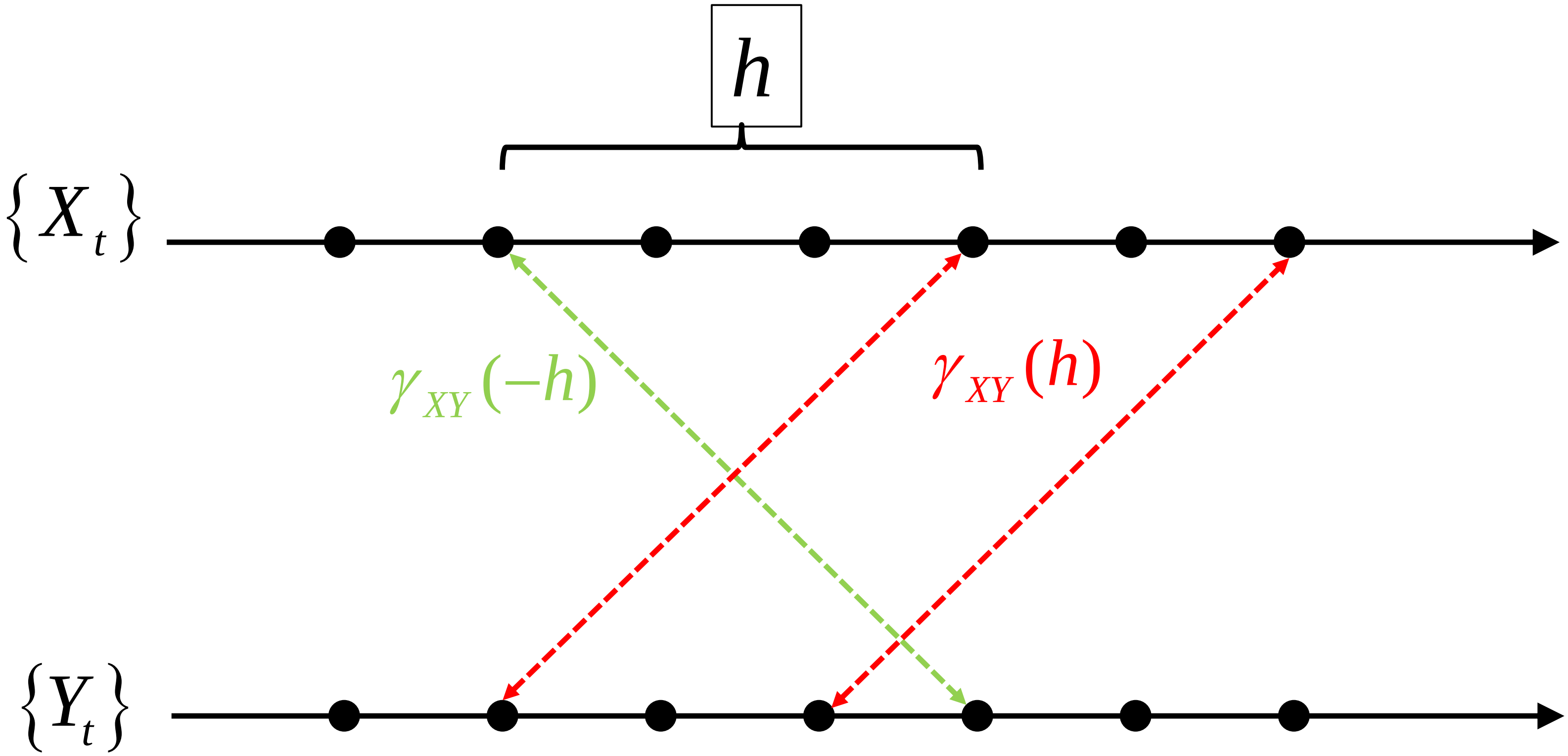
  - That's why we only look at $\gamma(h)$ for $h \geq 0$

- For jointly stationary $\{X_t, Y_t\}$, cross-covariance is *NOT* symmetric

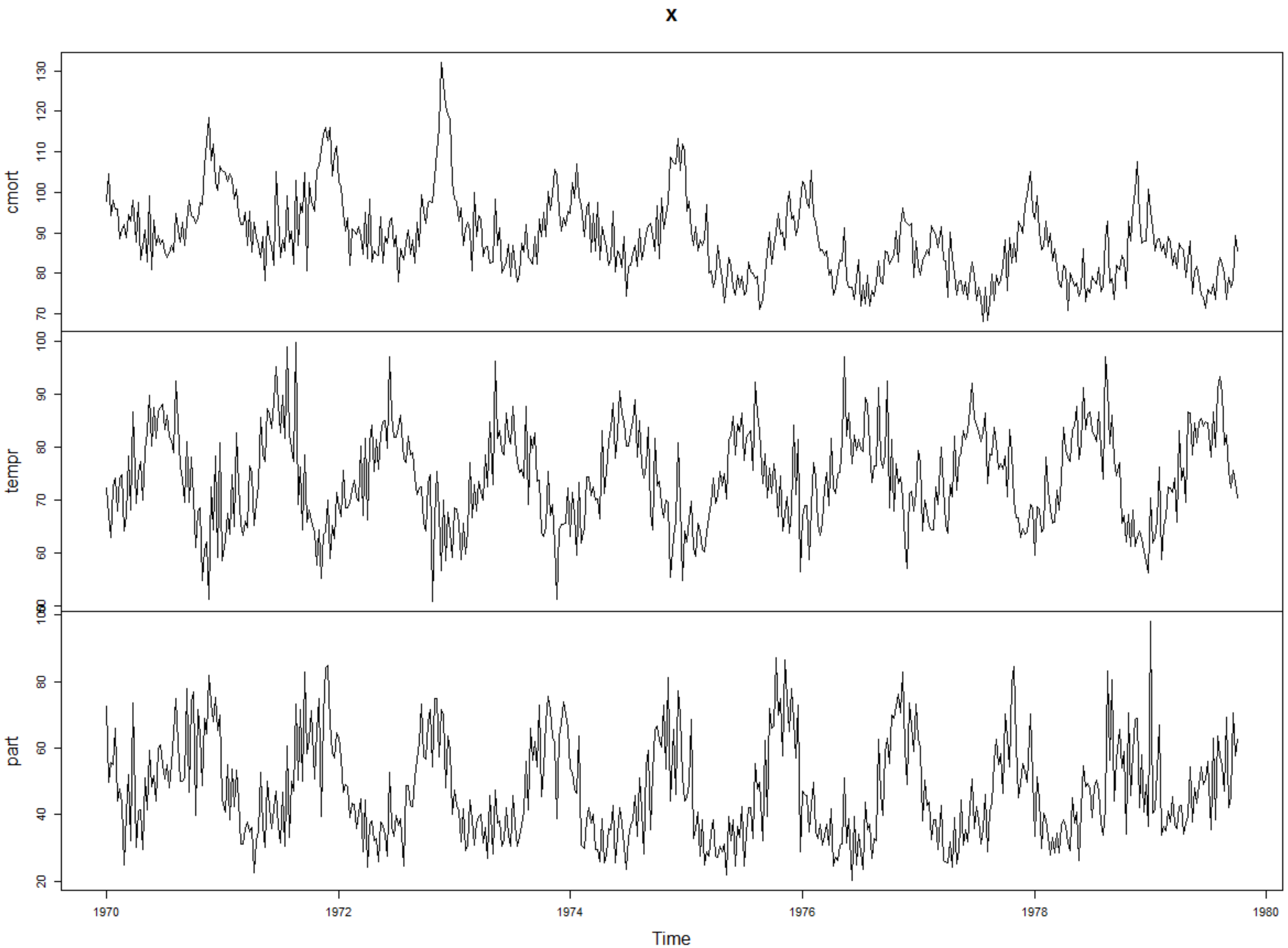
$$\gamma_{XY}(h) = \gamma_{XY}(t+h, t) = Cov[X_{t+h}, Y_t] \neq$$
$$\neq Cov[X_t, Y_{t+h}] = \gamma_{XY}(t-(t+h)) = \gamma_{XY}(-h)$$

  - That's why we need $\gamma_{XY}(h)$ for all $h=0,\pm1,\pm2,\ldots$
  - However, $\gamma_{\mathbf{XY}}(h) = \gamma_{\mathbf{YX}}(-h)$

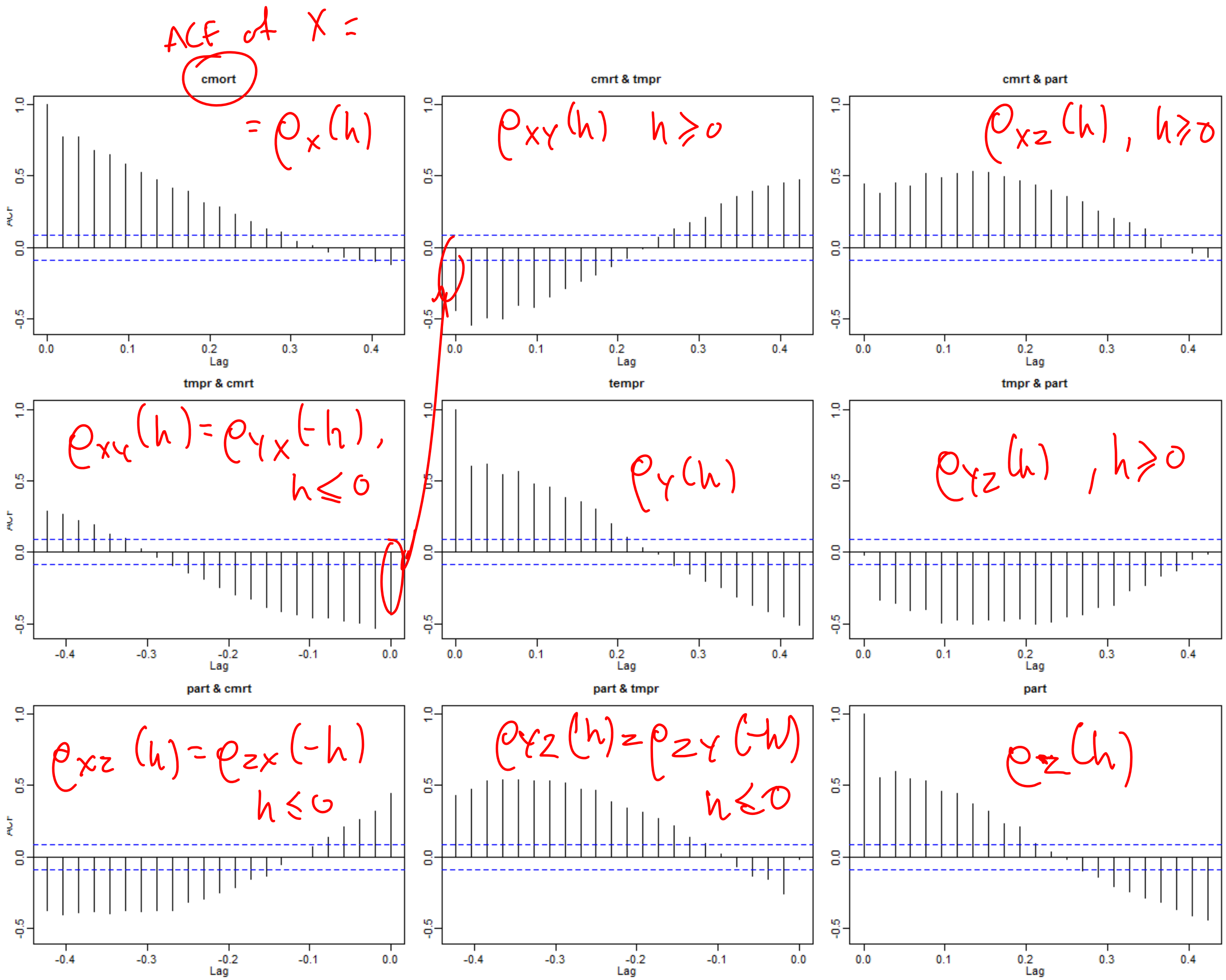
Auto- vs Cross-Covariance



Example



Example (Cont'd)



Example

- Consider 2D TS: $\begin{cases} X_t = W_t + W_{t-1} & \text{MA(1)} \\ Y_t = W_t - W_{t-1} & \text{MA(1)} \end{cases}$, where $W_t \sim WN(0, \sigma_W^2)$
  - Show that $\{X_t, Y_t\}$ is jointly stationary

First need to show that $\{X_t\}$ & $\{Y_t\}$ are individually stationary. But since they are both MA(1) $\Rightarrow$ they are both stationary with auto covariance function:

$$\gamma_X(h) = \begin{cases} 2\sigma_W^2, & h=0 \\ \sigma_W^2, & h=1 \\ 0, & h \geq 2 \end{cases}$$

$$\left(\begin{aligned} \gamma_X(1) &= Cov(X_{t+1}, X_t) = Cov(W_{t+1} + W_t, W_t + W_{t-1}) = \\ &= Cov(\cancel{W_{t+1}, W_t})^{0} + Cov(\cancel{W_{t+1}, W_{t-1}})^{0} + \underbrace{Cov(W_t, W_t)}_{=\sigma_W^2} + \\ &\quad + Cov(\cancel{W_t, W_{t-1}})^{0} \end{aligned} \right)$$

& similarly: $\gamma_Y(h) = \begin{cases} 2\sigma_W^2, & h=0 \\ -\sigma_W^2, & h=1 \\ 0, & h \geq 2 \end{cases}$

Calculate cross-covariance to show joint stationarity

$$\gamma_{XY}(X_t, Y_t) = Cov(X_t, Y_t) = Cov(W_t + W_{t-1}, W_t - W_{t-1}) = Cov(W_t, W_t) - \cancel{Cov(W_t, W_{t-1})}^{0} + \cancel{Cov(W_{t-1}, W_t)}^{0} - Cov(W_{t-1}, W_{t-1}) = \sigma_w^2 - \sigma_w^2 = 0 = \gamma_{XY}(0)$$

$$\gamma_{XY}(X_{t+1}, Y_t) = Cov(W_{t+1} + W_t, W_t - W_{t-1}) = \cancel{Cov(W_{t+1}, W_t)} - \cancel{Cov(W_{t+1}, W_{t-1})} + Cov(W_t, W_t) - \cancel{Cov(W_t, W_{t-1})} = \sigma_w^2 = \gamma_{XY}(h=1)$$

$$\gamma_{XY}(X_{t-1}, Y_t) = Cov(W_{t-1} + W_{t-2}, W_t - W_{t-1}) = \cancel{Cov(W_{t-1}, W_t)} - Cov(W_{t-1}, W_{t-1}) + \cancel{Cov(W_{t-2}, W_t)} - \cancel{Cov(W_{t-2}, W_{t-1})} = -\sigma_w^2 = \gamma_{XY}(h=-1)$$

$$\gamma_{XY}(X_{t+h}, Y_t) = \dots = 0, \ \forall |h| \geq 2 \Rightarrow$$

$$\Rightarrow \gamma_{XY}(h) = \begin{cases} 0, & h = 0 \\ +\sigma_w^2, & h = 1 \\ -\sigma_w^2, & h = -1 \\ 0, & |h| \geq 2 \end{cases} \Rightarrow \text{jointly stationary}$$

Example

$\{W_t\}$ uncorrelated w/ $\{X_t\}$

$\mu_x = 0$ & $\gamma_X(h)$ stationary

- Let $\{X_t\}$ be zero-mean, stationary TS & $W_t \sim WN(0,\sigma_w^2)$, and consider $Y_t = A \cdot X_{t-\ell} + W_t$
  - Show that $\{X_t, Y_t\}$ is jointly stationary

Know that $\{X_t\}$ is stationary $\Rightarrow$ need to show that $\{Y_t\}$ is also stationary (i.e. μ_Y is constant & $\gamma_Y(h)$ is stationary).

$$\mu_Y = \mathbb{E}[Y_t] = \mathbb{E}[A \cdot X_{t-\ell} + W_t] = A \cdot \mathbb{E}[\underbrace{X_{t-\ell}}_{=0}] + \mathbb{E}[\underbrace{W_t}_{=0}] = 0$$

$$\gamma_Y(s,t) = Cov(Y_s, Y_t) = Cov(A \cdot X_{s-\ell} + W_s, A \cdot X_{t-\ell} + W_t) =$$

$$= A^2 \cdot Cov(X_{s-\ell}, X_{t-\ell}) + A \cdot \underbrace{Cov(X_{s-\ell}, W_t)}_{0} + A \cdot \underbrace{Cov(W_s, X_{t-\ell})}_{0} + Cov(W_s, W_t)$$

$$= A^2 \gamma_X(s-\ell-t+\ell) + A \cdot 0 + A \cdot 0 + \sigma_w^2 \cdot 1_{(s=t)} =$$

$$= A^2 \cdot \gamma_X(h) + 1_{(h=0)} \cdot \sigma_w^2, \text{ where } h = |s-t|$$

$\Rightarrow \{Y_t\}$ is stationary

To show joint stationarity:

$$\gamma_{xy}(\overset{\rightarrow s=t+h}{t+h}, t) = Cov(X_{t+h}, Y_t) =$$

$$= Cov(X_{t+h}, A \cdot X_{t-\ell} + W_t) =$$

$$= A \cdot Cov(X_{t+h}, X_{t-\ell}) + \cancel{Cov(X_{t+h}, W_t)}^{\,0} =$$

$$= A\,\gamma_X(t+h-t+\ell) = A \cdot \gamma_X(h+\ell)$$

$\hookrightarrow$ function of $h = |s-t|$

Vector Auto-Regressive Model

- *Vector Auto-Regressive* (*VAR*) model

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{X}_{t-p} + \mathbf{W}_t$$

  - where

$$\mathbf{X}_t = \begin{bmatrix} X_{1,t} \\ \vdots \\ X_{k,t} \end{bmatrix} \,\&\, \mathbf{W}_t = \begin{bmatrix} W_{1,t} \\ \vdots \\ W_{k,t} \end{bmatrix}, \forall t \quad \mathbf{\Phi}_i = \begin{bmatrix} \varphi_{i:1,1} & \vdots & \varphi_{i:1,k} \\ \cdots & \ddots & \cdots \\ \varphi_{i:k,1} & \vdots & \varphi_{i:k,k} \end{bmatrix}, \forall i = 1, \ldots, p$$

$$\mathrm{Var}(\mathbf{W}_t) = \mathrm{Cov}(\mathbf{W}_t, \mathbf{W}_t) = \mathbf{\Sigma}_W = \begin{bmatrix} \sigma_{1,1}^2 & \vdots & \sigma_{1,k}^2 \\ \cdots & \ddots & \cdots \\ \sigma_{k,1}^2 & \vdots & \sigma_{k,k}^2 \end{bmatrix} \;\&\; \mathrm{Cov}(\mathbf{W}_t, \mathbf{W}_s) = \mathbf{0}, \forall s \neq t$$

Example

- 2D VAR(1) model:

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \mathbf{W}_t \Leftrightarrow \begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \varphi_{1;1,1} & \varphi_{1;1,2} \\ \varphi_{1;2,1} & \varphi_{1;2,2} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} W_{1,t} \\ W_{2,t} \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{Bmatrix} X_{1,t} = \varphi_{1;1,1} X_{1,t-1} + \varphi_{1;1,2} X_{2,t-1} + W_{1,t} \\ X_{2,t} = \varphi_{1;2,1} X_{1,t-1} + \varphi_{1;2,2} X_{2,t-1} + W_{2,t} \end{Bmatrix}$$

- where $\begin{cases} \mathbb{E}\left[W_{1,t}\right] = \mathbb{E}\left[W_{2,t}\right] = 0, \\ \mathbb{V}\left[W_{1,t}\right] = \sigma_{1,1}, \ \mathbb{V}\left[W_{2,t}\right] = \sigma_{2,2}, \\ \mathrm{Cov}(W_{1,t}, W_{2,t}) = \sigma_{1,2} = \sigma_{2,1} \end{cases}$

Example

- Find CCF of 2D $\{\mathbf{W}_t\} \sim \mathrm{WN}(\mathbf{0}, \mathbf{\Sigma}_W)$

$$\gamma_{W_1,W_2}(h) = Cov(W_{1,t+h}, W_{2,t}) = \begin{cases} \sigma_{1,2} = \sigma_{2,1}, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

$$\rho_{W_1,W_2}(h) = \frac{\gamma_{W_1,W_2}(h)}{\sqrt{\gamma_{W_1}(0) \cdot \gamma_{W_2}(0)}} = \begin{cases} \frac{\sigma_{1,2}}{\sqrt{\sigma_{1,1} \cdot \sigma_{2,2}}}, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

Example

- Fit VAR model with function `vars::VAR()`

```
VAR Estimation Results:
=======================

Estimated coefficients for equation cmort:
==========================================
Call:
cmort = cmort.l1 + tempr.l1 + part.l1 + const

   cmort.l1    tempr.l1     part.l1       const
 0.60149346 -0.30946101  0.07096225 54.94579126


Estimated coefficients for equation tempr:
==========================================
Call:
tempr = cmort.l1 + tempr.l1 + part.l1 + const

  cmort.l1   tempr.l1    part.l1      const
-0.1787076  0.5111817 -0.1412636 58.8456142


Estimated coefficients for equation part:
=========================================
Call:
part = cmort.l1 + tempr.l1 + part.l1 + const

   cmort.l1    tempr.l1     part.l1       const
-0.08082151 -0.45998718  0.57215788 61.58412402
```

→ VAR(1)

$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ & & \\ & & \end{bmatrix}$

VAR Model

- Consider VAR(p) model

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{X}_{t-p} + \mathbf{W}_t, \ \{\mathbf{W}_t\} \sim \mathrm{WN}(\mathbf{0}, \mathbf{\Sigma}_W)$$

- Model can be written as Wold process

$$\mathbf{X}_t = \mathbf{W}_t + \mathbf{\Psi}_1 \mathbf{W}_{t-1} + \mathbf{\Psi}_2 \mathbf{W}_{t-2} + \cdots = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \mathbf{W}_{t-j}$$

  - where $\mathbf{\Psi}$-matrices satisfy:

$$\mathbf{\Psi}_k = \sum_{j=0}^{\min(k,p)} \mathbf{\Psi}_{k-j} \mathbf{\Phi}_j \ \& \ \mathbf{\Psi}_0 = \mathbf{I}$$

Example

- Find Wold representation of VAR(1) model

$$\mathbf{X}_t = \mathbf{\Phi X}_{t-1} + \mathbf{W}_t,\ \{\mathbf{W}_t\} \sim \mathrm{WN}(\mathbf{0}, \mathbf{\Sigma}_W)$$

$$\begin{aligned}
\underline{X}_t &= \underline{\Phi}\,\underline{X}_{t-1} + \underline{W}_t \\
&= \underline{\Phi}\cdot(\underline{\Phi}\cdot\underline{X}_{t-2} + \underline{W}_{t-1}) + \underline{W}_t \\
&= \underline{\Phi}^2 X_{t-2} + \underline{\Phi}\,\underline{W}_{t-1} + W_t \\
&= \underline{\Phi}^2\cdot(\underline{\Phi}\,X_{t-3} + \underline{W}_{t-2}) + \underline{\Phi}\,\underline{W}_{t-1} + W_t \\
&\ \vdots \\
&= \sum_{j=0}^{\infty} \underline{\Phi}^j \cdot \underline{W}_{t-j}, \quad \text{where } \underline{\Phi}^0 = \underline{I}
\end{aligned}$$

Example (cont'd)

- Find stationary variance-covariance matrix of VAR(1) model

$$\mathbb{E}[\underline{X}_t] = \mathbb{E}\left[\sum_{j=0}^{\infty} \underline{\underline{\Phi}}^j \cdot \underline{W}_{t-j}\right] = \sum_{j=0}^{\infty} \underline{\underline{\Phi}}^j \underbrace{\mathbb{E}[\underline{W}_{t-j}]}_{=\underline{0}} = \underline{0}$$

$$\mathrm{Var}[\underline{X}_t] = \mathrm{Var}\left[\sum_{j=0}^{\infty} \underline{\underline{\Phi}}^j \cdot \underline{W}_{t-j}\right] = \sum_{j=0}^{\infty} \underline{\underline{\Phi}}^j \cdot \underbrace{\mathrm{Var}[\underline{W}_{t-j}]}_{=\Sigma_w} \left(\underline{\underline{\Phi}}^j\right)^T =$$

$$= \sum_{j=0}^{\infty} \underline{\underline{\Phi}}^j \, \underline{\underline{\Sigma}}_w \cdot \left(\underline{\underline{\Phi}}^j\right)^T$$

VAR Model

- Any VAR(p) can be expressed as special VAR(1) model:

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{X}_{t-p} + \mathbf{W}_t \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} \mathbf{X}_t \\ \vdots \\ \mathbf{X}_{t-p+1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \cdots & \mathbf{\Phi}_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p} \end{bmatrix} + \begin{bmatrix} \mathbf{W}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

VAR Model

- A VAR(p) model is causal (stationary) if

$$\det\left(\mathbf{I} - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p\right) \neq 0, \ \forall \, |z| \leq 1$$

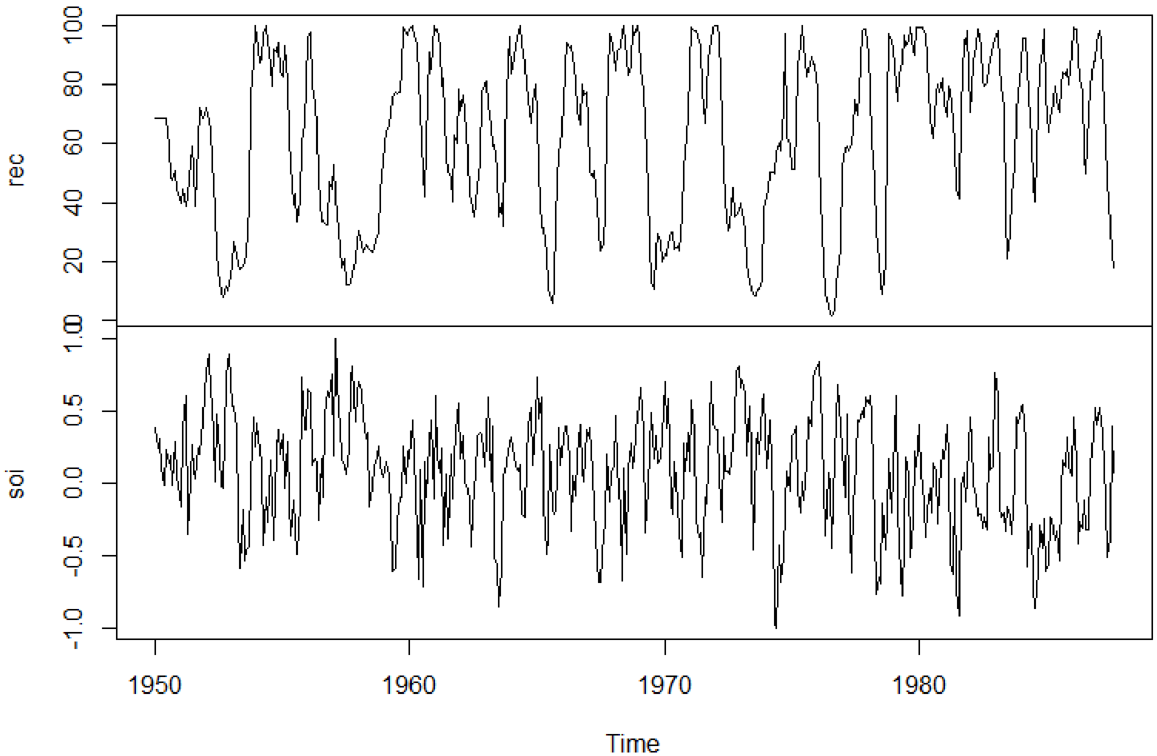
  - For VAR(1) ⇔ eigen-values of $\mathbf{\Phi}$ are all <1
  - Generally ⇔ eigen-values of $\begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \cdots & \mathbf{\Phi}_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \end{bmatrix}$ are all <1

VAR Estimation

- Function `VAR()` in package `vars` fits VAR(p) model
  - E.g. `VAR(X, p=2)`
- Model selection using AIC/BIC through `VARselect()` function
  - E.g. `VARselect(X, lag.max=15)`
  - Returns "AIC", "BIC" & other criteria

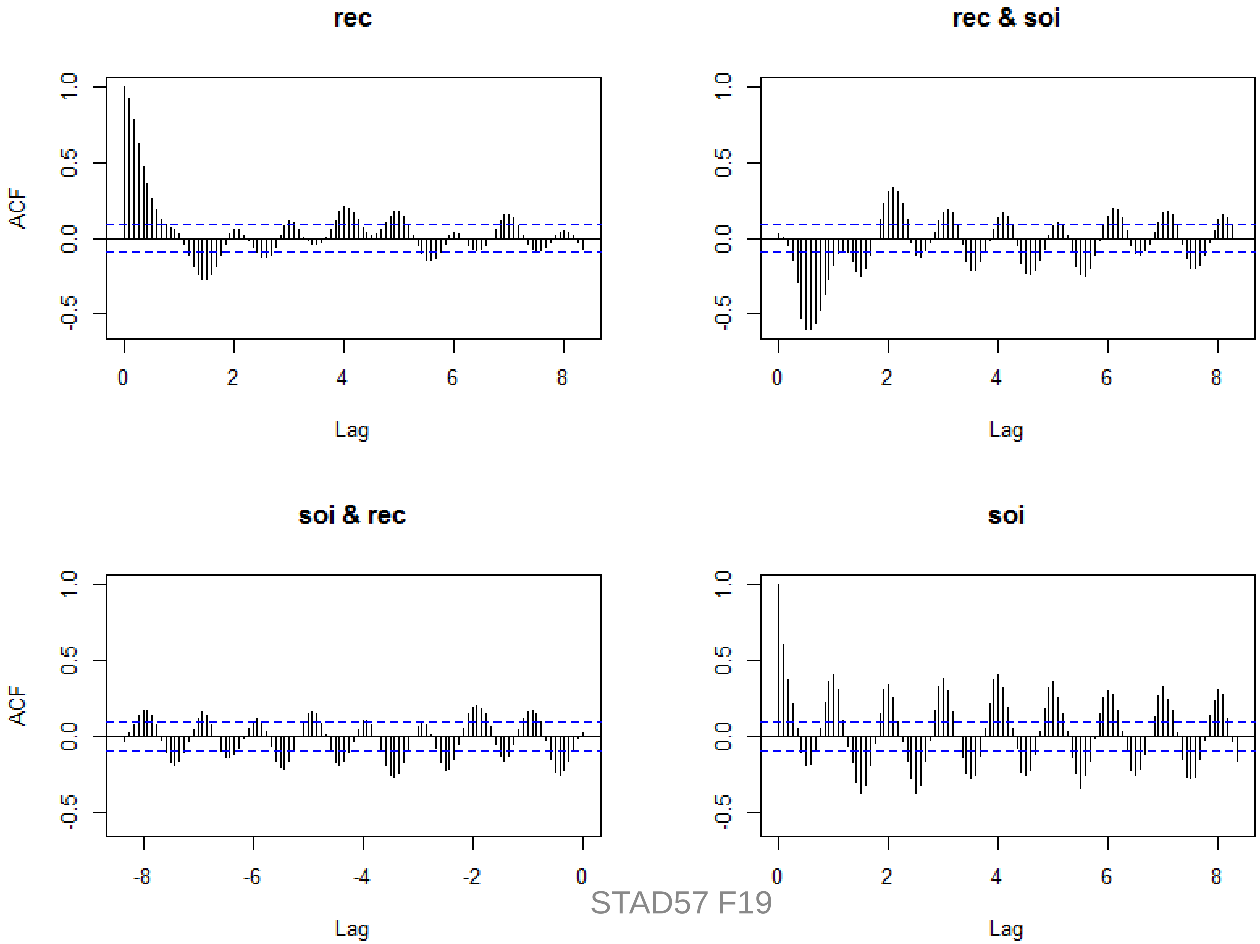
Example

- Monthly data: Southern Oscillation Index (soi) & Pacific Ocean # fish (rec)



Example (cont'd)

- ACF / CCF



Example (cont'd)

- VAR model selection

BIC = SC = "Schwartz Criterion"

```
> VARselect(X,lag.max=20)
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
    15      8      7     15

$criteria
                1        2        3        4        5        6        7        8        9       10
AIC(n)   2.402970 2.140677 2.112579 2.027653 1.650150 1.542299 1.488852 1.470325 1.464954 1.449333
HQ(n)    2.425237 2.177789 2.164536 2.094455 1.731797 1.638791 1.600188 1.596507 1.605981 1.605205
SC(n)    2.459377 2.234689 2.244197 2.196875 1.856978 1.786731 1.770889 1.789968 1.822202 1.844186
FPE(n) 11.055963 8.505208 8.269589 7.596325 5.207876 4.675494 4.432249 4.351002 4.327834 4.260922
              11       12       13       14       15       16       17       18       19       20
AIC(n) 1.444060 1.446837 1.445788 1.408018 1.402028 1.411512 1.407134 1.414192 1.424015 1.439889
HQ(n)  1.614776 1.632399 1.646194 1.623269 1.632124 1.656453 1.666920 1.688823 1.713491 1.744209
SC(n)  1.876517 1.916900 1.953455 1.953290 1.984905 2.031994 2.065221 2.109884 2.157312 2.210791
FPE(n) 4.238715 4.250746 4.246571 4.089489 4.065427 4.104585 4.087123 4.116601 4.157834 4.225032
```

  - Can include deterministic seasonality w/ option `VARselect(…,season=s)`

Example (cont'd)

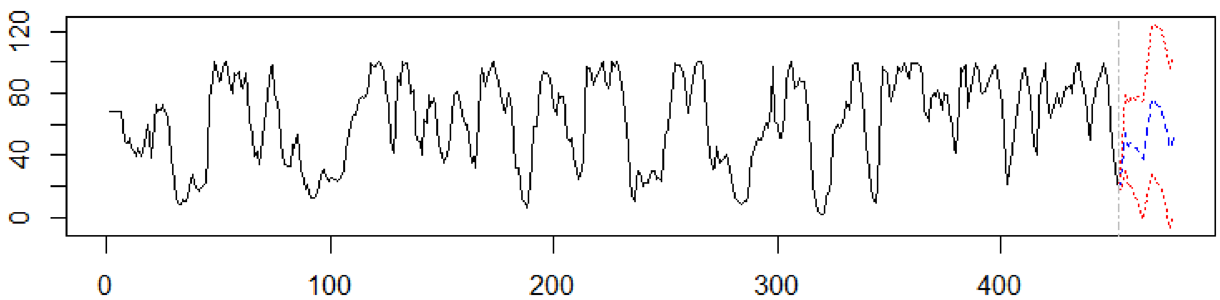
- VAR(15) estimation

```
> VAR(X,15)

VAR Estimation Results:
=======================

Estimated coefficients for equation rec:
========================================
Call:
rec = rec.l1 + soi.l1 + rec.l2 + soi.l2 + rec.l3 + soi.l3 + rec.l4 + soi.l4 + rec.l5 + soi.l5 + rec.l6 +
soi.l6 + rec.l7 + soi.l7 + rec.l8 + soi.l8 + rec.l9 + soi.l9 + rec.l10 + soi.l10 + rec.l11 + soi.l11 + r
ec.l12 + soi.l12 + rec.l13 + soi.l13 + rec.l14 + soi.l14 + rec.l15 + soi.l15 + const

       rec.l1        soi.l1        rec.l2        soi.l2        rec.l3        soi.l3        rec.l4
  1.207691796   1.174911739  -0.393793460   0.433512840   0.012686405  -1.393827810  -0.149158384
       soi.l4        rec.l5        soi.l5        rec.l6        soi.l6        rec.l7        soi.l7
  0.019510371   0.199015432 -21.480745005   0.017758950   9.115855299  -0.209012841  -1.888180940
       rec.l8        soi.l8        rec.l9        soi.l9       rec.l10       soi.l10       rec.l11
  0.195335535  -2.403517603  -0.121119906  -2.774691973  -0.011415651  -0.093660952  -0.003666367
      soi.l11       rec.l12       soi.l12       rec.l13       soi.l13       rec.l14       soi.l14
 -0.711908788   0.060769283  -4.103027176  -0.039959051   2.923913383  -0.041878162  -1.594892567
      rec.l15       soi.l15         const
  0.001724510  -0.826725126  19.000121895
```

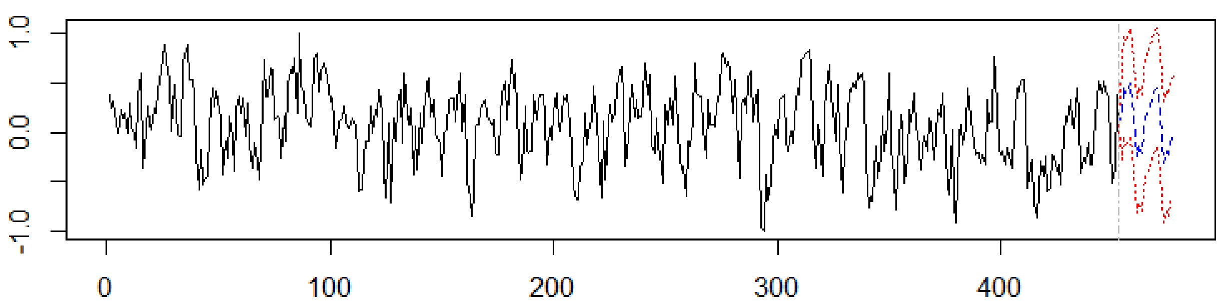
Example (cont'd)

- Predictions w/ `predict()` function

Forecast of series rec



Forecast of series soi



Granger Causality

- Consider 2D VAR(1) model

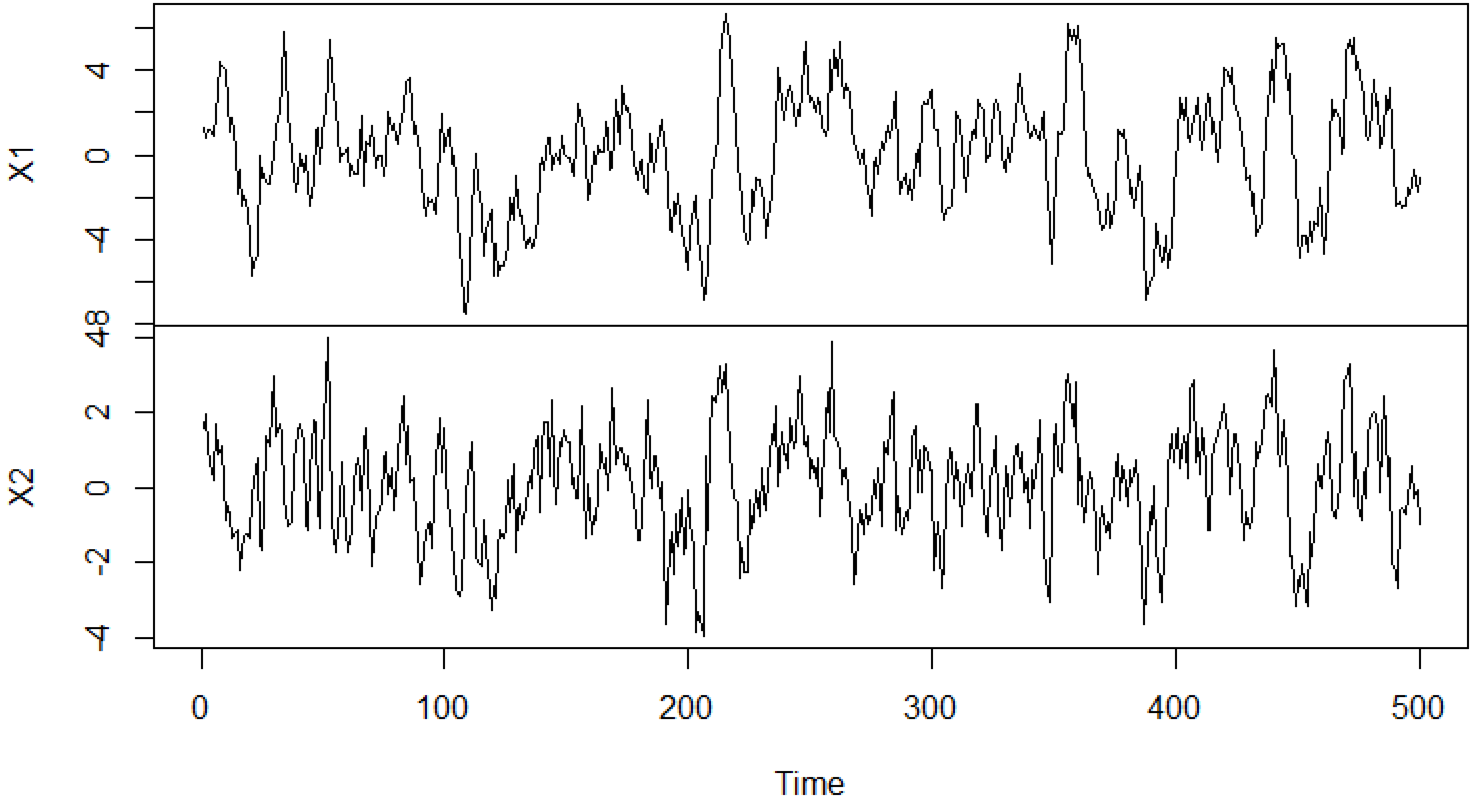
$$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} \\ 0 & \varphi_{2,2} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} W_{1,t} \\ W_{2,t} \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \left\{ \begin{array}{c} X_{1,t} = \varphi_{1,1} X_{1,t-1} + \varphi_{1,2} X_{1,t-2} + W_{1,t} \\ X_{2,t} = \varphi_{2,2} X_{2,t-1} + W_{2,t} \end{array} \right\}$$

  - Coordinate $X_{1,t}$ depends on *both* $X_{1,t-1}$ & $X_{2,t-1}$, but coordinate $X_{2,t}$ depends on $X_{2,t-1}$ *only*
  - Nevertheless, both coordinates can be cross-correlated at different lags

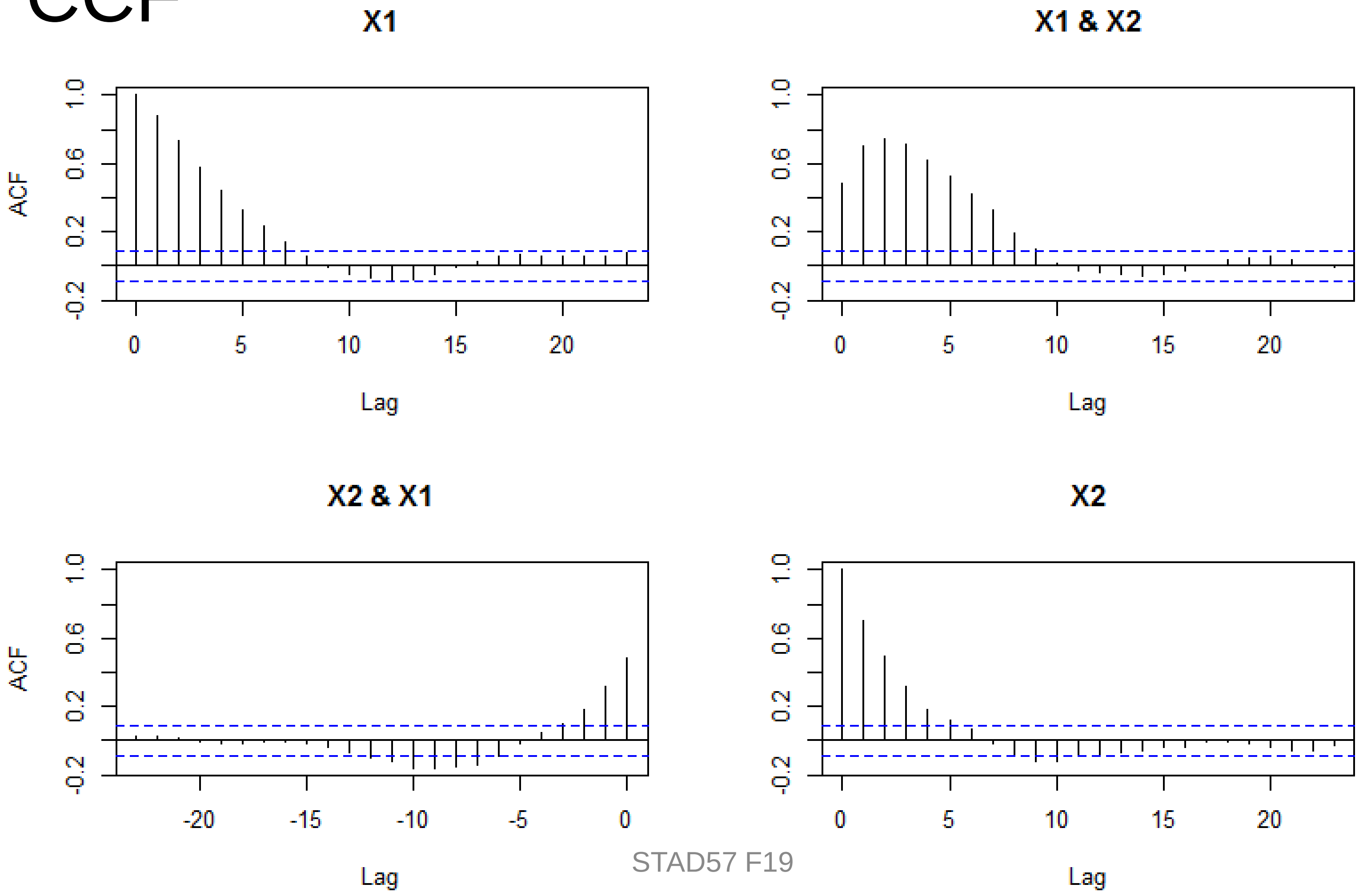
Example

- Simulated series from $\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} .7 & .7 \\ 0 & .7 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} W_{1,t} \\ W_{2,t} \end{bmatrix}$



Example (cont'd)

- CCF



Granger Causality

- TS $\{Y_t\}$ is said to *Granger-cause* TS $\{X_t\}$ if past of $\{Y_t\}$ helps in predicting $\{X_t\}$ beyond using past of $\{X_t\}$ alone
  - Clive Granger, 2003 Nobel prize in Economics
- In terms of VAR(p) model, Granger-causality implies certain structure of zero-coefficients in the dynamic equation

Example

- For VAR(1) $\begin{bmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{bmatrix} = \begin{bmatrix} \varphi_{1;1,1} & \varphi_{1;1,2} & 0 \\ 0 & \varphi_{1;2,2} & \varphi_{1;2,3} \\ 0 & 0 & \varphi_{1;3,3} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \\ X_{3,t-1} \end{bmatrix} + \begin{bmatrix} W_{1,t} \\ W_{2,t} \\ W_{3,t} \end{bmatrix}$

  find which variable Granger-causes which

For predicting $\{X_1\}$: $\{X_2\}$ Granger causes $\{X_1\}$ (b/c of $\varphi_{1,2}$)

$\{X_3\}$ does not Granger-cause $\{X_1\}$ (b/c given (X_1, X_2), then X_3 is not used for prediction)

For predicting $\{X_2\}$: $\{X_1\}$ does not Granger cause $\{X_2\}$

$\{X_3\}$ Granger causes $\{X_2\}$

For prediction $\{X_3\}$: neither $\{X_1\}$ nor $\{X_2\}$ Granger causes $\{X_3\}$

Granger Causality

- Granger-causality based on VAR(p) model

$$\texttt{causality(}\underbrace{\texttt{VAR.fit}}_{\texttt{VAR() output}}\texttt{, cause=}\underbrace{\texttt{"X1"}}_{\text{cause TS name}}\texttt{)}$$

- R output:

```
> causality(out, cause='X1')
$Granger

 Granger causality H0: X1 do not Granger-cause X2

data:  VAR object out
F-Test = 0.52434, df1 = 1, df2 = 992, p-value = 0.4692


$Instant

 H0: No instantaneous causality between: X1 and X2

data:  VAR object out
Chi-squared = 0.27869, df = 1, p-value = 0.5976
```

```
> causality(out, cause='X2')
$Granger

 Granger causality H0: X2 do not Granger-cause X1

data:  VAR object out
F-Test = 366.14, df1 = 1, df2 = 992, p-value < 2.2e-16


$Instant

 H0: No instantaneous causality between: X2 and X1

data:  VAR object out
Chi-squared = 0.27869, df = 1, p-value = 0.5976
```

Impulse Response Function

- Want to know how changes in one coordinate affect others
  - Assume: $$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} W_{1,t} \\ W_{2,t} \end{bmatrix}$$
  - Let $$\begin{bmatrix} X_{1,0} \\ X_{2,0} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \ \& \ \begin{bmatrix} W_{1,1} \\ W_{2,1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
  - How does 1unit change in $W_{1,\mathrm{t}}$ propagate through time?

Example

$$\begin{bmatrix} x_{1,1} \\ x_{2,1} \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} X_{1,0}=0 \\ X_{2,0}=0 \end{bmatrix} + \begin{bmatrix} w_{1,1}=1 \\ w_{2,1}=0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} X_{1,2} \\ X_{2,2} \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} X_{1,1}=1 \\ X_{2,1}=0 \end{bmatrix} + \begin{bmatrix} W_{1,2}=0 \\ W_{2,2}=0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} X_{1,3} \\ X_{2,3} \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} X_{1,2} \\ X_{2,2} \end{bmatrix}} + \underline{0} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} X_{1,1}=1 \\ X_{2,1}=0 \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} X_{1,t+1} \\ X_{2,t+1} \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix}^t \cdot \begin{bmatrix} x_{t,1}=1 \\ x_{2,1}=0 \end{bmatrix} \Rightarrow \left(\text{for VAR(1)}: \quad \underline{X}_{t+1} = \underline{\Phi}^t \cdot \underline{X}_1 \right)$$

Impulse Response Function

- Use causal (Wold) representation of VAR(p) model to trace effect of impulse

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{X}_{t-p} \Leftrightarrow$$

$$\mathbf{X}_t = \mathbf{W}_t + \mathbf{\Psi}_1 \mathbf{W}_{t-1} + \mathbf{\Psi}_2 \mathbf{W}_{t-2} + \cdots = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \mathbf{W}_{t-j}$$

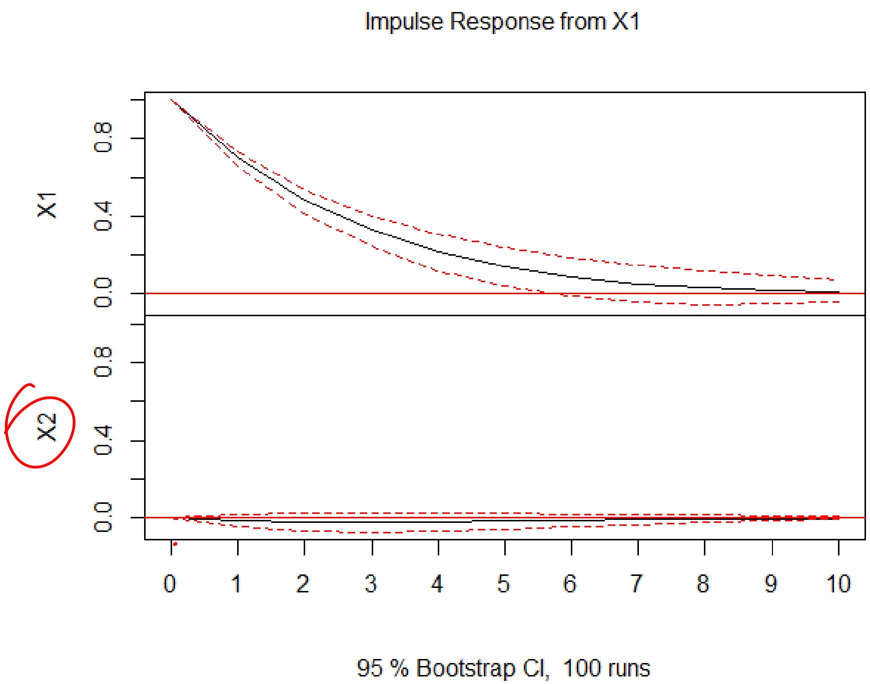
where $\mathbf{\Psi}_k = \sum_{j=0}^{\min(k,p)} \mathbf{\Psi}_{k-j} \mathbf{\Phi}_j$ & $\mathbf{\Psi}_0 = \mathbf{I}$

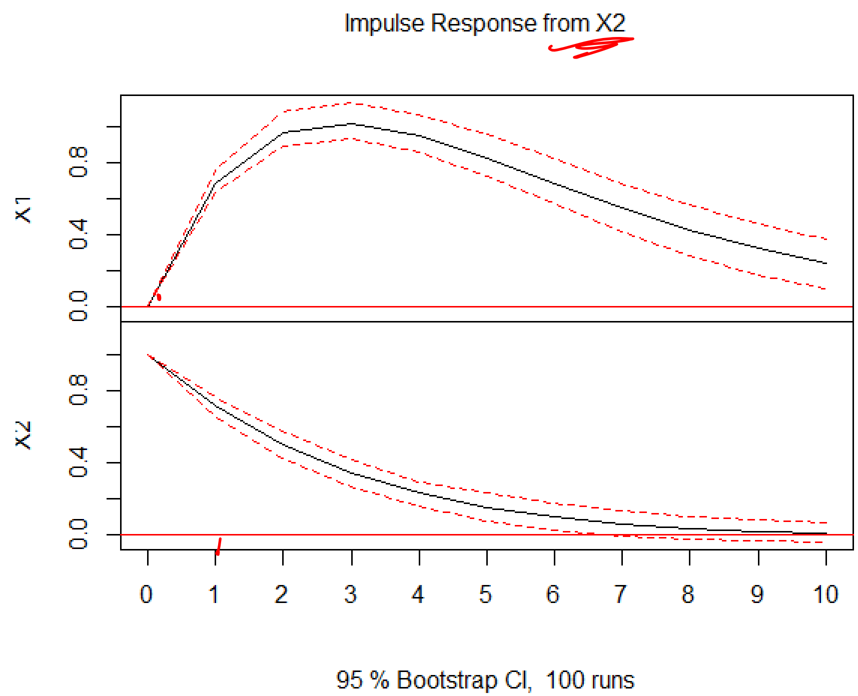
- Impulse Response Function (IRF) is given by components of $\mathbf{\Psi}$-matrices

Example

$W_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$W_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



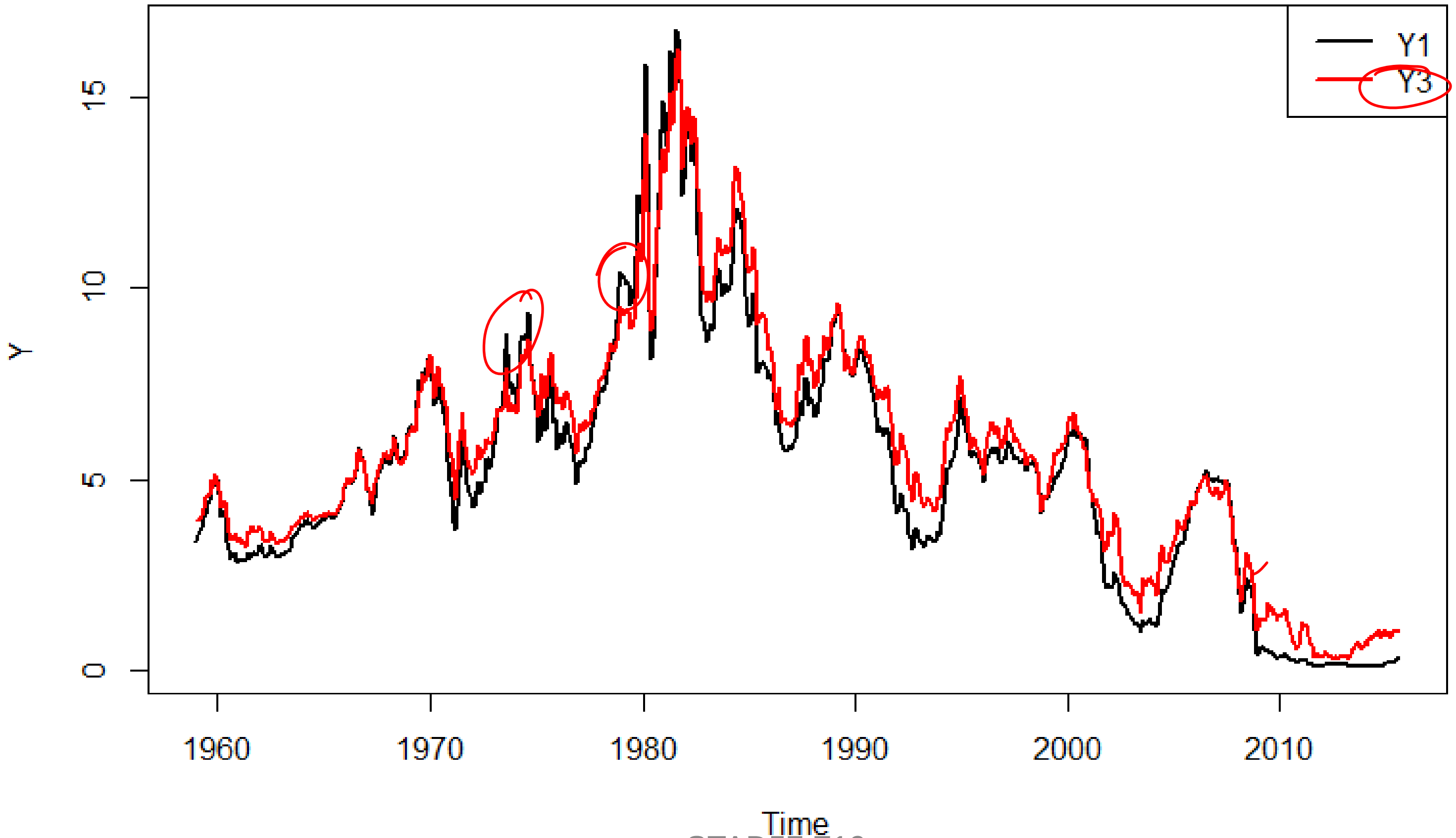


Cointegration

- Set of TS called *cointegrated* if :
  - Individual TS are *integrated*
    - e.g. follow $I(1)$~random walk
  - Some *linear combination* thereof is *stationary*
- E.g. Term-structure of interest rates
  - Consider yield rates of Gov't issued bonds with different maturities: e.g. 1yr vs 3yr
  - Interest levels fluctuate like a random walk, but rates for different maturities are close

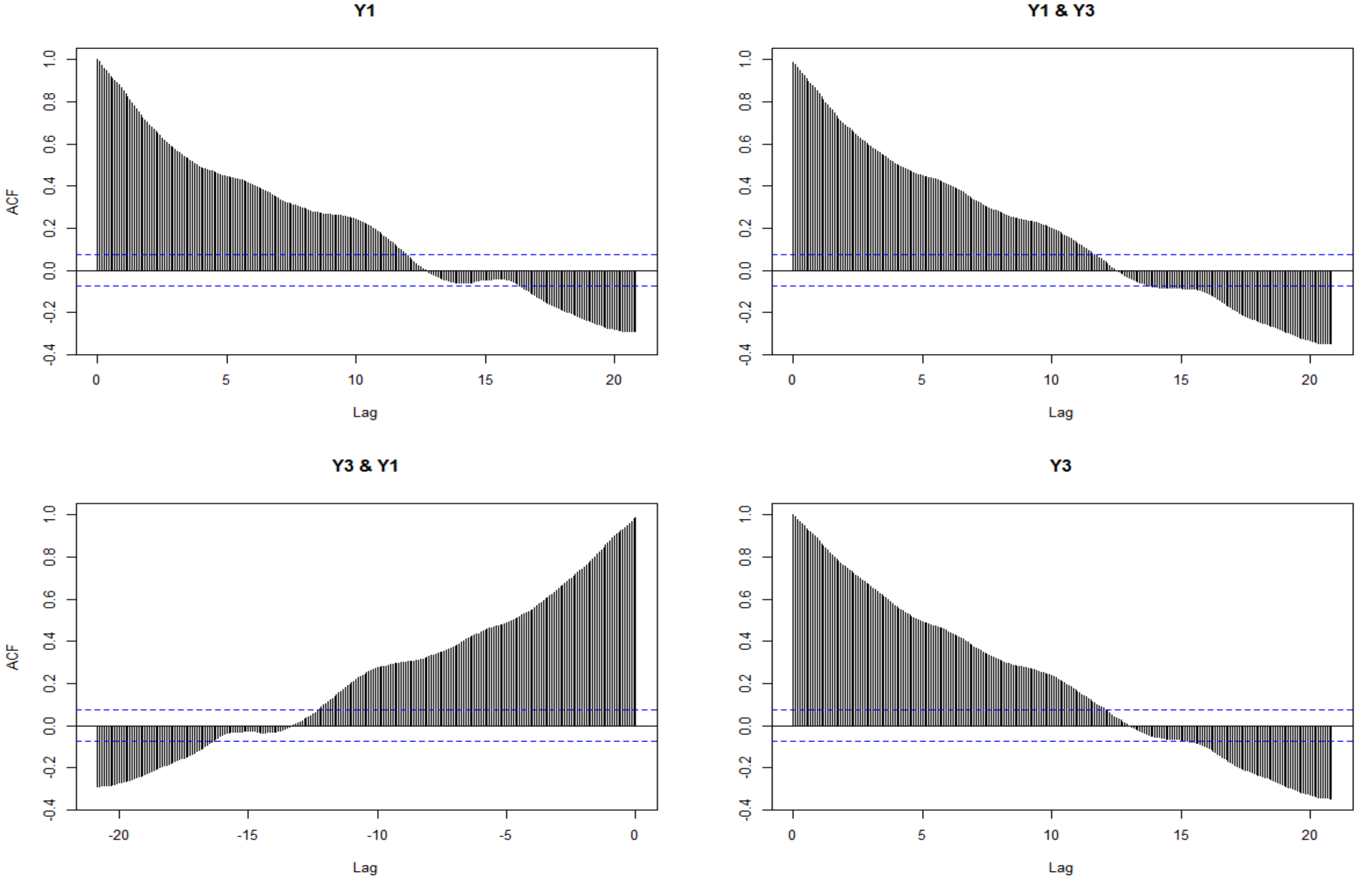
Example

- 1- & 3-year US Gov't bond yield rates



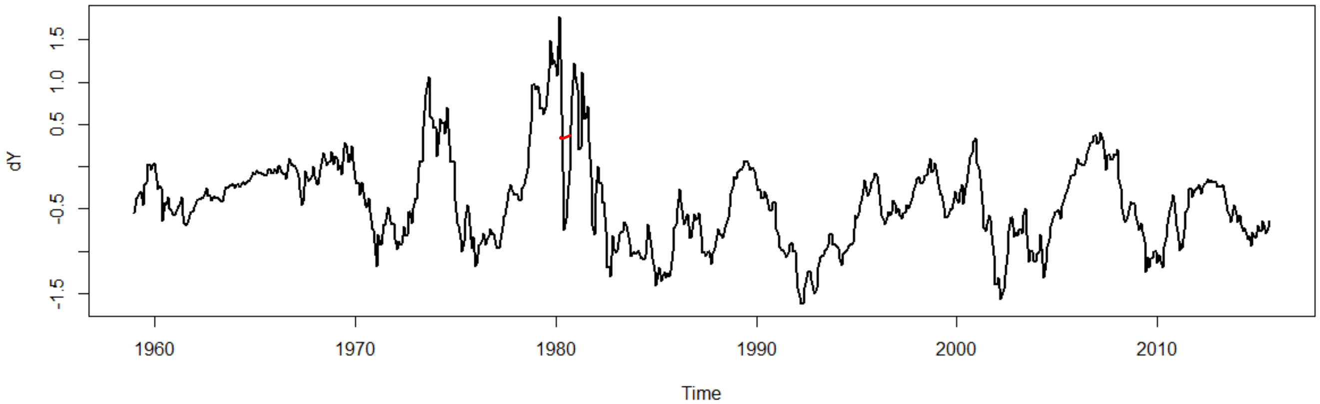
Example (cont'd)

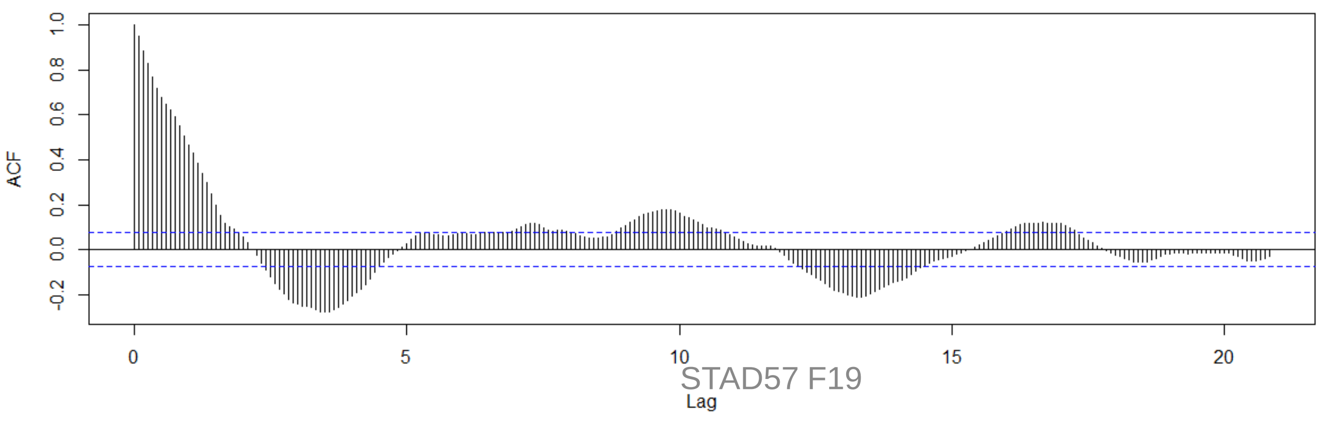
- ACF/CCF



Example (cont'd)

- Difference Y1-Y3





Cointegration

- If you *know stationary relation*, just test it for stationarity (w/ unit root tests)
  - E.g. Augmented Dickey Fuller (ADF) test
- If you *don't know stationary relation*, need to estimate; two approaches
  - Engle-Granger two-step process
    1. Regression to estimate stationary relation
    2. Perform unit root test on residuals
  - Johansen test, using VAR models

Example (cont'd)

- Unit root tests (ADF) on Y1, Y3 & Y1-Y3

```
> adf.test(Y[,"Y1"])

 Augmented Dickey-Fuller Test

data:  Y[, "Y1"]
Dickey-Fuller = -2.0811, Lag order = 8, p-value = 0.544
alternative hypothesis: stationary

> adf.test(Y[,"Y3"])

 Augmented Dickey-Fuller Test

data:  Y[, "Y3"]
Dickey-Fuller = -1.8658, Lag order = 8, p-value = 0.6351
alternative hypothesis: stationary

> adf.test(dY)

 Augmented Dickey-Fuller Test

data:  dY
Dickey-Fuller = -3.9589, Lag order = 8, p-value = 0.01105
alternative hypothesis: stationary
```

$H_0: \varphi_1 = 1$

$H_1: \varphi_1 < 1$

; don't reject H_0 ⇒ integrated

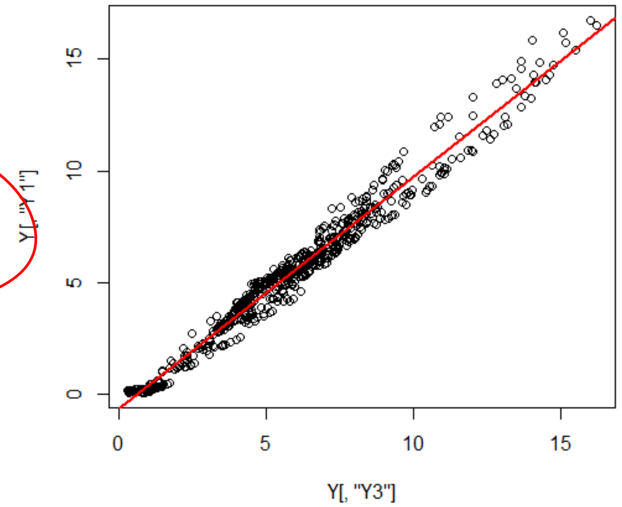
$\leq 5\%$

Example (cont'd)

- Engle-Granger

```
> (out=lm(Y1~Y3,data=Y))

Call:
lm(formula = Y1 ~ Y3, data = Y)

Coefficients:
(Intercept)           Y3
    -0.6348       1.0363

> dY.est=residuals(out)
> adf.test(dY.est)

 Augmented Dickey-Fuller Test

data:  dY.est
Dickey-Fuller = -3.9004, Lag order = 8, p-value = 0.01398
alternative hypothesis: stationary
```

$Y_1 - 1.0363\,Y_3 + .6348$



Spurious Regression

- Consider *independent* random walks $\{W_t, V_t\}$
  - When you regress $W_t = \beta_0 + \beta V_t + e_t,\ t = 1, \ldots, n$ you are NOT guaranteed that $\hat{\beta} \to 0$ as the sample size $n \to \infty$ (i.e. not consistent)!!!

- Effect called *spurious* (fake) regression
  - Results of random walk (integrated series) regressions are NOT reliable

Cointegration & VAR

- Consider VAR (p)

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{X}_{t-p} + \mathbf{W}_t$$

  - Model is stable (causal) if

$$\det\left(\mathbf{I} - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p\right) \neq 0, \ \forall\, |z| \leq 1$$

  - If there is a *unit root*, then all or some of the coordinates of $\mathbf{X}_t$ are $I(1)$
- If model is cointegrated, some linear combination of $\mathbf{X}_t$ are $I(0)$

Cointegration & VAR

- Write VAR model as *Vector Error Correction Model* (*VECM*)

$$\Delta \mathbf{X}_t = \mathbf{\Lambda} \mathbf{X}_{t-1} + \mathbf{\Lambda}_1 \Delta \mathbf{X}_{t-1} + \cdots + \mathbf{\Lambda}_{p-1} \Delta \mathbf{X}_{t-p+1} + \mathbf{W}_t \Rightarrow$$

- where $\begin{cases} \mathbf{\Lambda} = \mathbf{\Phi}_1 + \cdots + \mathbf{\Phi}_p - \mathbf{I} \\ \mathbf{\Lambda}_i = -(\mathbf{\Phi}_{i+1} + \cdots + \mathbf{\Phi}_p) = \sum_{k=i+1}^{p} \mathbf{\Phi}_k \end{cases}$

$X_t \sim I(1)$. Show VECM for $\Delta X_t \iff \underline{\text{VAR}}$ for X_t

$\rightarrow$ has unit root

$$\Delta X_t = \underbrace{\Lambda X_{t-1}}_{} + \Lambda_1 \Delta X_{t-1} + \dots + \Lambda_{p-1} \Delta X_{t-p+1} + W_t$$

$$= (\Phi_1 + \cancel{\Phi_2} + \cancel{\Phi_3} + \dots + \cancel{\Phi_p} - I) \cdot X_{t-1}$$

$$- (\Phi_2 + \cancel{\Phi_3} + \dots + \cancel{\Phi_p}) \cdot (\cancel{X_{t-1}} - X_{t-2})$$

$$- (\Phi_3 + \dots + \cancel{\Phi_p}) \cdot (\cancel{X_{t-2}} - X_{t-3})$$

$$\vdots$$

$$- \Phi_p \cdot (\cancel{X_{t-p+1}} - X_{t-p}) + W_t$$

$$\Rightarrow X_t - \cancel{X_{t-1}} = \Phi_1 X_{t-1} - \cancel{X_{t-1}} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + W_t$$

$$X_t = \sum_{j=1}^{p} \Phi_j X_{t-j} + W_t \Rightarrow \text{VAR}(p)$$

Cointegration & VAR

- For the VECM

$$\Delta \mathbf{X}_t = \mathbf{\Lambda X}_{t-1} + \mathbf{\Lambda}_1 \Delta \mathbf{X}_{t-1} + \cdots + \mathbf{\Lambda}_{p-1} \Delta \mathbf{X}_{t-p+1} + \mathbf{W}_t, \quad \mathbf{X} \in \mathbb{R}^d$$

  - $\{\Delta \mathbf{X}_t\}$ is $I(0)$, but $\{\mathbf{X}_t\}$ is $I(1)$
  - Thus, term $\mathbf{\Lambda X}_{t-1}$ must also be $I(0) \rightarrow \mathbf{\Lambda}$ must contain cointegration relation(s)
  - Since $\det\left(\mathbf{I} - \mathbf{\Phi}_1 - \cdots - \mathbf{\Phi}_p\right) = \det\left(\mathbf{\Lambda}\right) = 0$ from unit root of $\{\mathbf{X}\}$, matrix $\mathbf{\Lambda}$ has *reduced rank* ($r<d$), i.e can be written as $\underset{(d\times d)}{\mathbf{\Lambda}} = \underset{(d\times r)}{\boldsymbol{\alpha}} \underset{(r\times d)}{\boldsymbol{\beta}^\top}$, where β defines cointegrating relations

Example

- For $\begin{cases} \nabla X_{1,t} = \varphi_1 \left(X_{1,t-1} - \lambda X_{2,t-1} \right) + \varepsilon_{1,t} \\ \nabla X_{2,t} = \varphi_2 \left(X_{1,t-1} - \lambda X_{2,t-1} \right) + \varepsilon_{2,t} \end{cases}$, show that

$Y_t = X_{1,t} - \lambda X_{2,t}$ follows AR(1) process

$$\nabla \underline{X}_t = \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} [1 \ -\lambda] \underline{X}_{t-1}$$

Take 1st line & substract $\lambda \times$ 2nd line:

$$\nabla X_{1,t} - \lambda \nabla X_{2,t} = \varphi_1 (X_{1,t-1} - \lambda X_{2,t-1}) + \varepsilon_1 - \lambda \varphi_2 (X_{1,t-1} - \lambda X_{2,t-1})$$

$$\Rightarrow (\varphi_1 - \lambda \varphi_2) \cdot X_{1,t-1} - \lambda (\varphi_1 - \lambda \varphi_2) X_{2,t-1} + \varepsilon_{1,t} - \lambda \varepsilon_{2,t}$$

$$\Rightarrow (\varphi_1 - \lambda \varphi_2) \cdot (X_{1,t-1} - \lambda X_{2,t-1})$$

$$\nabla X_{1,t} - \lambda \nabla X_{2,t} = (\varphi_1 - \lambda\varphi_2)(X_{1,t-1} - \lambda X_{2,t-1}) + \underbrace{\varepsilon_{1,t} - \lambda\varepsilon_{2,t}}_{\varepsilon_t}$$

$$\Rightarrow \underbrace{(X_{1,t} - \lambda X_{2,t})}_{Y_t} - \underbrace{(X_{1,t-1} - \lambda X_{2,t-1})}_{Y_{t-1}} = $$

$$= (\varphi_1 - \lambda\varphi_2)\underbrace{(X_{1,t-1} - \lambda X_{2,t-1})}_{Y_{t-1}} + \varepsilon_t$$

$$\Rightarrow Y_t = \underbrace{(1 + \varphi_1 - \lambda\varphi_2)}_{= \varphi} Y_{t-1} + \varepsilon_t$$

$Y_t = \varphi Y_{t-1} + \varepsilon_t$, which is stationary
$|\varphi| < 1$

Cointegration & VAR

- Johansen procedure:
  - Specify and estimate VAR(p) model for $\{\mathbf{X}_t\}$
  - Construct Likelihood Ratio(LR) tests for the rank of $\mathbf{\Lambda}$, to determine number of cointegrating vectors
    - If necessary, impose normalization and identifying restrictions on the cointegrating vectors.
  - Given cointegrating vectors, estimate resulting VECM by maximum likelihood

Example

- Find # of cointegrating vectors

```
> out=ca.jo(Y, ecdet="const", K=3)
> summary(out)

######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegratio
n

Eigenvalues (lambda):
[1] 3.930817e-02 3.907291e-03 4.336809e-18

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  2.65  7.52  9.24 12.97
r = 0  | 27.15 13.75 15.67 20.20

Eigenvectors, normalised to first column:
(These are the cointegration relations)

              Y1.l3     Y3.l3  constant
Y1.l3     1.0000000  1.000000  1.000000
Y3.l3    -1.0159606 -2.429144 -1.729806
constant  0.5080471  7.914550 18.018859
```

Example (cont'd)

- Fit VECM model

$$\Lambda = \alpha \cdot \beta^T = \begin{bmatrix} -.0244 \\ .03177 \end{bmatrix} \begin{bmatrix} 1 & -1.0159 \end{bmatrix}$$

```
> cajorls(out,r=1)
$rlm

Call:
lm(formula = substitute(form1), data = data.mat)

Coefficients:
         Y1.d      Y3.d
ect1    -0.02448   0.03177
Y1.dl1   0.20721   0.05366
Y3.dl1   0.30138   0.35856
Y1.dl2  -0.16093  -0.02903
Y3.dl2  -0.09866  -0.20990


$beta
              ect1
Y1.l3     1.0000000
Y3.l3    -1.0159606
constant  0.5080471
```

α

Λ_1

Λ_2

β

$$\Delta X_t = \mu + \Lambda X_{t-1} + \Lambda_1 \Delta X_{t-1} + \Lambda_2 \cdot \Delta X_{t-2}$$