Multivariate Time Series (Chapter 7)

7.1 Examples

Bivariate Time Series : $\{\mathbf{X}_t = (X_{t1}, X_{t2})'\}$.

Mean vector : $\boldsymbol{\mu}_t = \begin{bmatrix} EX_{t1} \\ EX_{t2} \end{bmatrix}$

Covariance matrices : $\Gamma(t+h, t) := \mathrm{Cov}(\mathbf{X}_{t+h}, \mathbf{X}_t)$

$\{\mathbf{X}_t = (X_{t1}, X_{t2})'\}$ is (**weakly**) **stationary** if $\boldsymbol{\mu}_t$ and $\Gamma(t+h, t)$ do not depend on t. In this case, write

$$\boldsymbol{\mu} = \begin{bmatrix} EX_{t1} \\ EX_{t2} \end{bmatrix} \quad \text{and} \quad \Gamma(h) := \mathrm{Cov}(\mathbf{X}_{t+h}, \mathbf{X}_t)$$

Properties of

$$\Gamma(h) = \begin{bmatrix} \gamma_{11}(h) & \gamma_{12}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) \end{bmatrix}, \quad \gamma_{ij}(h) = \mathrm{Cov}(X_{t+h,i}, X_{t,j}).$$

(i) $\Gamma(h) = \Gamma'(-h)$,

(ii) $|\gamma_{12}(h)| \le [\gamma_{11}(0)\,\gamma_{22}(0)]^{1/2}$,

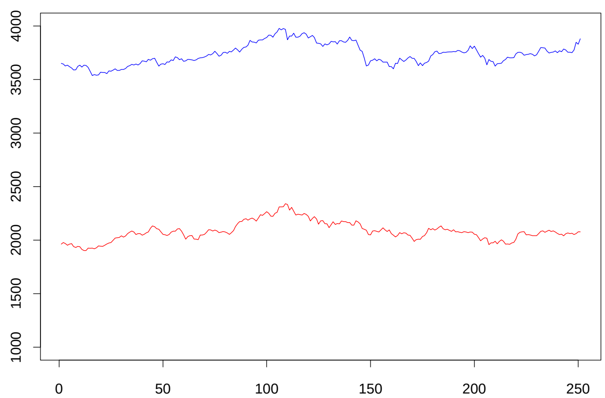
(iii) $\gamma_{11}(h)$ and $\gamma_{22}(h)$ are the ACVFs of $\{X_{t1}\}$ and $\{X_{t2}\}$

(iv) $\Gamma(h)$ is a non-negative definite matrix sequence.

Ex 7.1.1 (Dow-Jones, All-ordinaries: DJA02.TSM).

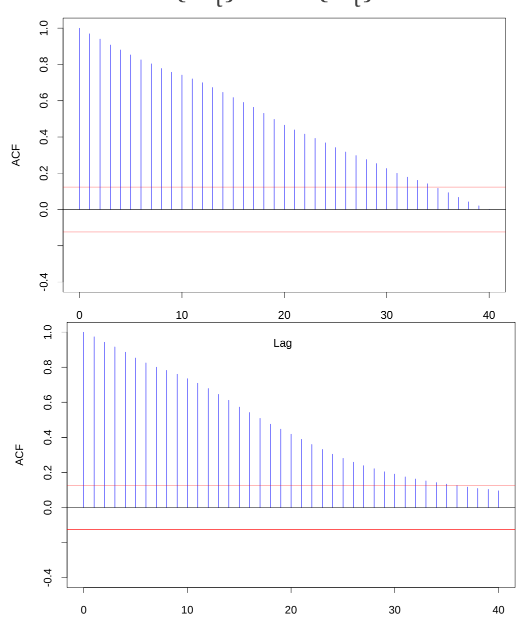
D_j = closing values of Dow-Jones, j= 0, . . ., 250.

A_j = closing values Australian All-ordinaries, j= 0, . . ., 250.



3

Sample ACFs of $\{D_t\}$ and $\{A_t\}$:



4

Efficient market hypothesis suggests modelling this data as random walks.

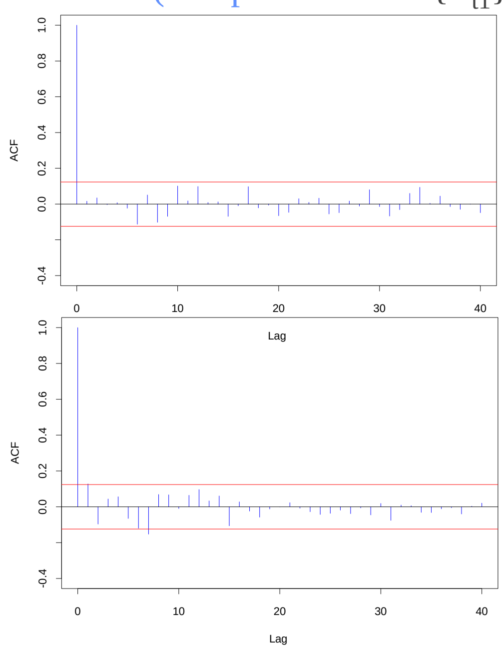
Transform data to percentage relative price changes (DJAOPC2.TSM) given by:

$$X_{t1} = 100\,(D_t - D_{t-1})/\,D_{t-1}, \quad t = 1, \ldots, 250,$$

$$X_{t2} = 100\,(A_t - A_{t-1})/\,A_{t-1}, \quad t = 1, \ldots, 250.$$

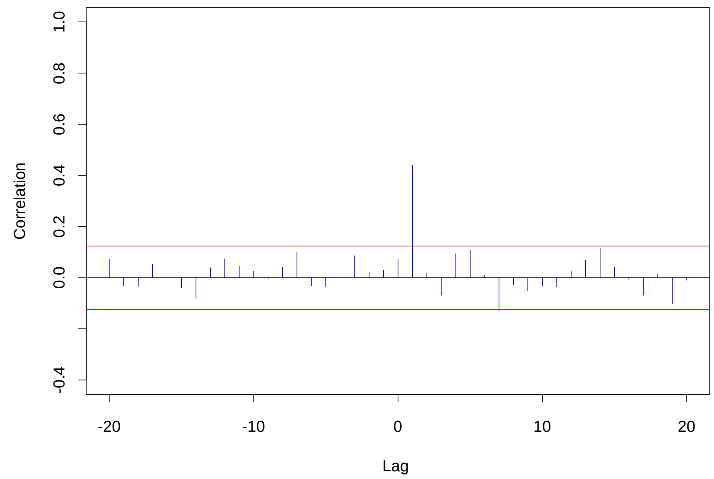
5

Figs 7.2 & 7.3 (Sample ACF's of $\{X_{t1}\}$ and $\{X_{t2}\}$).



6

Fig 7.4 Sample cross-correlations $\hat{\rho}_{21}(h)$ between $X_{t-h,1}$ and X_{t2}.



Graph indicates correlation between $X_{t-1,1}$ and X_{t2}.

7.6 Modelling with Multivariate AR

Multivariate white noise: $\{\mathbf{Z}_t\} \sim \mathrm{WN}(\mathbf{0},\Sigma)$ means

- $E\mathbf{Z}_t = \mathbf{0}.$
- $E(\mathbf{Z}_t\mathbf{Z}_s') = \begin{cases} \Sigma, & \text{if } s=t, \\ 0, & \text{otherwise.} \end{cases}$

Multivariate AR(p) Process:

$\{\mathbf{X}_t\}$ is an **AR(p) process** if $\{\mathbf{X}_t\}$ is stationary and

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \mathbf{Z}_t\,,\ \{\mathbf{Z}_t\} \sim \mathrm{WN}(\mathbf{0},\Sigma)$$

Causality :

An **AR(p) process** $\{\mathbf{X}_t\}$ is **causal**, or a **causal function** of $\{\mathbf{Z}_t\}$ if there exist matrices $\{\Psi_j\}$ such that

$$\mathbf{X}_t = \sum_{j=0}^{\infty} \Psi_j \mathbf{Z}_{t-j} .$$

Result: Causality is equivalent to the condition

$$\det(\Phi(z)) \neq 0 \quad \text{for all } |z| \le 1,$$

where

$$\Phi(z) = I - \Phi_1 z - \cdots - \Phi_p z^p.$$

9

Yule-Walker Equations :

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \mathbf{Z}_t , \quad \{\mathbf{Z}_t\} \sim \mathrm{WN}(\mathbf{0},\Sigma).$$

Post multiply with $\mathbf{X}_{t-j}'$ and take expectations to get

$$\Sigma = \Gamma(0) - \sum_{j=1}^{p} \Phi_j \Gamma(j),$$

$$\Gamma(i) = \sum_{j=1}^{p} \Phi_j \Gamma(i-j), \quad i = 1, \ldots, p.$$

Given $\Gamma(0), \ldots, \Gamma(p)$, these equations can be solved for $\Phi_1, \ldots, \Phi_p$ and Σ.

10

Yule-Walker Estimates:

Replace $\Gamma(i)$ with $\hat{\Gamma}(i)$, $i=1, \ldots, p$ in the Y-W eqns and solve for $\Phi_1, \ldots, \Phi_p$ and Σ.

- Method uses Whittle's algorithm.
- These estimates are computed in ITSM2000

Order Selection:

Choose p to minimize the AICC statistic

$$\text{AICC} = -2 \ln L(\Phi_1, \ldots, \Phi_p, \Sigma) + \frac{2(pm^2+1)nm}{nm-pm^2-2}$$

11

Ex 7.6.1 (Dow-Jones , All ordinaries indices).

- In ITSM, open new project and select multivariate for project type.
- Select the data file DJAOPC2.TSM
- Enter 2 for the number of columns
- Click on the blue button labeled Y-W (Program finds the Y-W AR(p), p=0, ..., 20, with minimum AICC.)

12

Minimum AICC Y-W AR model is

$$\begin{bmatrix} X_{t1} \\ X_{t2} \end{bmatrix} - \begin{bmatrix} .0262 \\ .0268 \end{bmatrix} = \begin{bmatrix} .0146 & .0177 \\ .6493 & .0958 \end{bmatrix} \left(\begin{bmatrix} X_{t-1,1} \\ X_{t-1,2} \end{bmatrix} - \begin{bmatrix} .0262 \\ .0268 \end{bmatrix} \right) + \begin{bmatrix} Z_{t1} \\ Z_{t2} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} .3478 & .0334 \\ .0334 & .6319 \end{bmatrix}$$

AICC =1048.6

13

7.6.2 Forecasting Multivariate AR

One-step predictors :

Since

$$\mathbf{X}_{n+1} = \Phi_1 \mathbf{X}_n + \cdots + \Phi_p \mathbf{X}_{n+1-p} + \mathbf{Z}_{n+1},$$

it follows that for $n > p$.

$$\hat{\mathbf{X}}_{n+1} = P_n \mathbf{X}_{n+1} = \Phi_1 \mathbf{X}_n + \cdots + \Phi_p \mathbf{X}_{n+1-p}$$

h-step predictors :

Found from the recursions

$$P_n \mathbf{X}_{n+h} = \Phi_1 P_n \mathbf{X}_{n+h-1} + \cdots + \Phi_p P_n \mathbf{X}_{n+h-p}$$

14

Forecasting All-ordinaries:

1. Using the multivariate AR(1) model.

$$\hat{X}_{251,2} = .0268 + .6493\,(X_{250,1} - .0262) + .0958(X_{250,2} - .0268) = .869$$

with MSE

$$E(X_{251,2} - \hat{X}_{251,2})^2 = .6319\,.$$

15

2. Using a univariate AR(1) model.

$$X_{t,2} = .0273 + .1180\,X_{t-1,2} + Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0,.7604).$$

Best predictor is

$$\hat{X}_{251,2} = .0273 + .11180 X_{250,2} = .6016$$

with MSE

$$E(X_{251,2} - \hat{X}_{251,2})^2 = .\,7604\,.$$

16

Summary of one-step prediction MSE

MSE	Model
.6319	multivariate AR(1)model
.7678	univariate AR(2)
.7911	white noise

Comparison of one-step predictors on next 40 observations ($\mathbf{X}_{251}, \ldots, \mathbf{X}_{290}$).

Observed MSE	Model
.3931	multivariate AR(1)model
.3947	univariate AR(2)
.4699	white noise

Remark: Last 40 observations less variable than the first 250.

17

Ex 7.6.2 (Sales with a Leading Indicator; LS2.DAT).

- In ITSM, open new project and select multivariate for project type.
- Select the data file LS2.TSM
- Enter 2 for the number of columns
- Difference the data at lag 1 using Transfer
- Click on the blue button labeled Y-W
  (Program finds the Y-W AR(p), p=0, ..., 20, with minimum AICC.)
- Can also try Burg estimates (use BRG button).

18

Minimum AICC Y-W AR model has order 5 with

$$\hat{\Phi}_1 = \begin{bmatrix} -.517 & .024 \\ -.019 & -.051 \end{bmatrix}, \quad \hat{\Phi}_2 = \begin{bmatrix} -.192 & -.018 \\ .047 & .250 \end{bmatrix}$$

$$\hat{\Phi}_3 = \begin{bmatrix} -.073 & .010 \\ 4.678 & .207 \end{bmatrix}, \quad \hat{\Phi}_4 = \begin{bmatrix} -.032 & .009 \\ 3.664 & .004 \end{bmatrix}$$

$$\hat{\Phi}_5 = \begin{bmatrix} .022 & .011 \\ 1.300 & .029 \end{bmatrix}, \quad \hat{\Sigma} = \begin{bmatrix} .076 & -.003 \\ -.003 & .095 \end{bmatrix}$$

Note: Upper right component of each coefficient matrix is nearly 0 $\Rightarrow$ X_{t1} can be modelled independent of X_{t2}

19

Model for X_{t1}:

$$X_{t1} = (1-.474B)U_t, \quad \{U_t\} \sim \mathrm{WN}(0,.0779)$$

Model for X_{t2}. Setting all the small values in the bottom rows of the estimated coefficients to 0, we obtain

$$X_{t2} = .250\, X_{t-2,2} + .207\, X_{t-3,2} + 4.678\, X_{t-3,1} + 3.664\, X_{t-4,1} + 1.300\, X_{t-5,1} + Z_{t2}$$

or equivalently,

$$X_{t2} = \frac{4.678\, B^3(1+.783\, B+.278\, B^2)}{1-.250\, B^2-.207\, B^3} X_{t1} + \frac{Z_{t1}}{1-.250\, B^2-.207\, B^3}$$

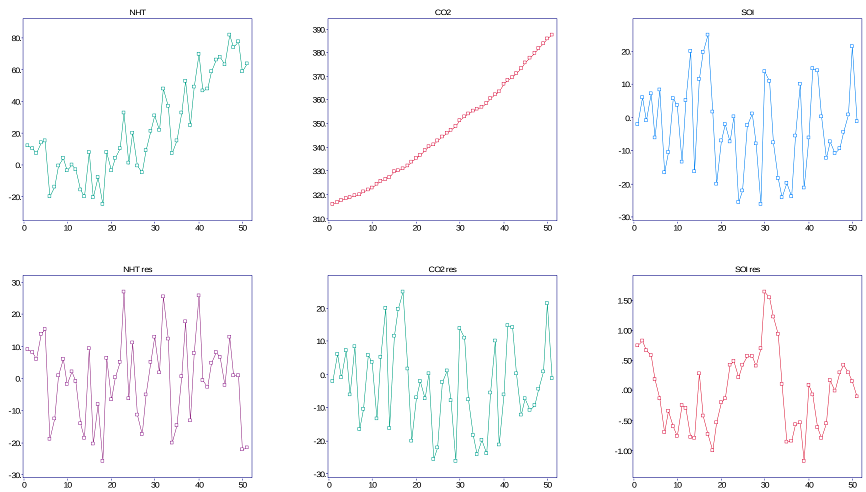
(Transfer function model (see Section 10.1).)

20

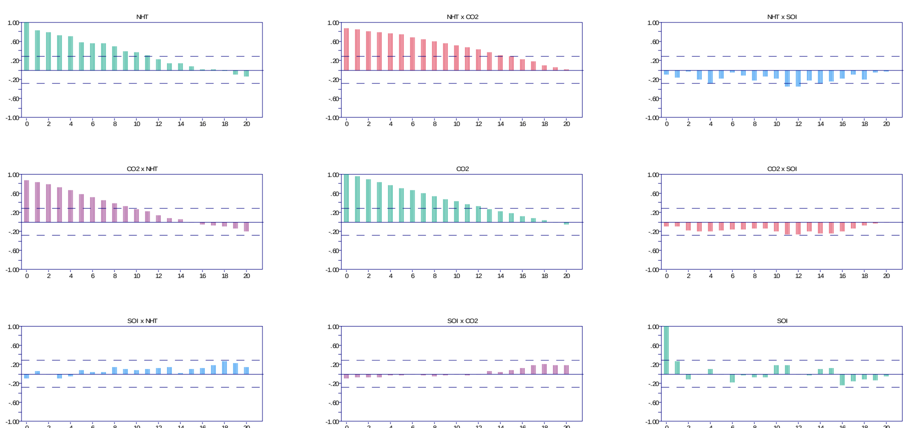
Northern Hemisphere temperature data example:(see Tunnicliffe Wilson `10)

Data: years t= 1959 to 2009 (annual)

- Average temp in northern hemisphere: T_t
- CO_2 in Mauna Loa: C_t
- Southern Oscillation Index: S_t

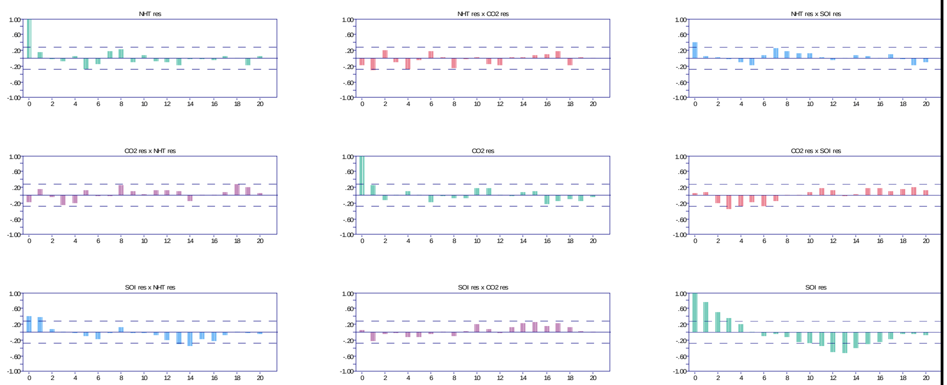


Cross-correlation function for original data



22

Cross-correlation function for residuals from quadratic fit



23

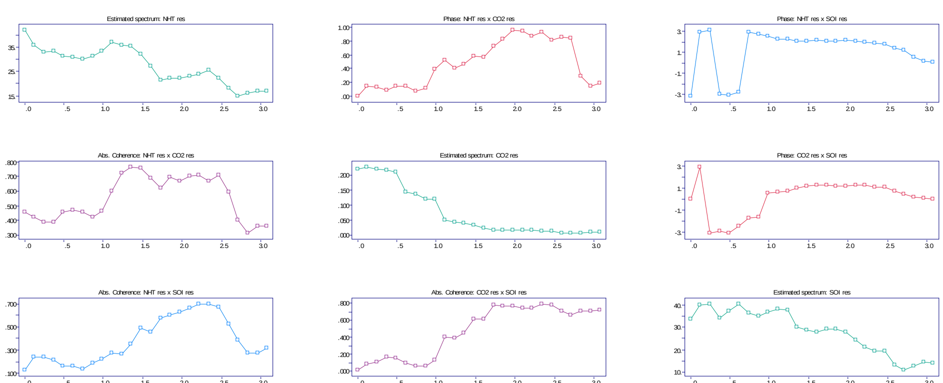
Testing for 0 Coherency $\kappa(\omega)$:

Reject $\kappa(\omega) = 0$ at level .05 if $Y > F_{\{1-\alpha\}}(2{,}4m)$, where

$$Y = 2m\frac{|\hat{\kappa}(\omega)|^2}{1-|\hat{\kappa}(\omega)|^2}$$

and a constant coefficient filter of length 2m+1 is used in the smoothing.

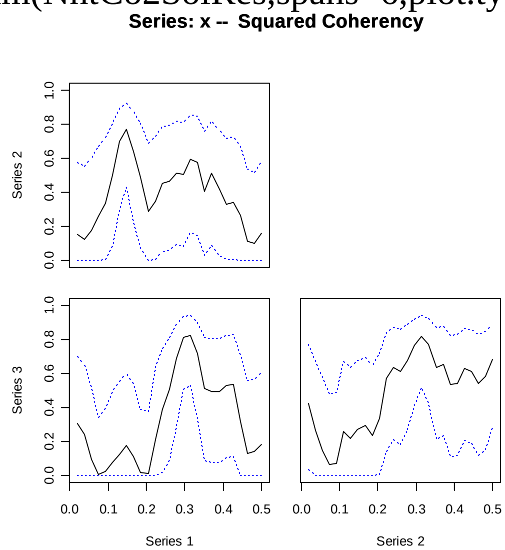
For m=6, reject if $|\hat{\kappa}(\omega)| > .470 \Rightarrow$ correlation at freq ω



24

Coherency plot in R:

spectrum(NhtCo2SoiRes,spans=6,plot.type="coh")



25

Interpretation:

• Slope of phase $\phi_{\{1,2\}}(\omega)$ corresponds to the lag of the second component behind the first. That is, if $\phi_{\{1,2\}}(\omega)$ is d, then $X_{t,1}$ will lead $X_{t,2}$ by d lags. So in the previous graphs, we see that the slope of $\phi_{\{1,2\}}(\omega)$ and $\phi_{\{1,3\}}(\omega)$ is slightly negative for large frequencies ω which suggests temperature follows CO2 and SOI.

• The coherency $\kappa(\omega)$ is a measure of correlation between series at frequency ω. In this case, the partial coherency for the original data (not trend corrected) gives similar results as the coherency for the detrended data.

• These methods are often used for constructing vector AR models using graphical models.

26

Best fitting VAR model to detrended data has order 2:

Optimal value of p = 2

PHI(0)

.095850
-.008256
-2.398619

PHI(1)

.104591	5.561032	-.398659
.000391	1.071852	-.016209
.048084	9.027895	.294766

PHI(2)

.070975	-4.431574	.429041
-.006298	-.247972	.013592
-.061732	-10.778030	-.055049

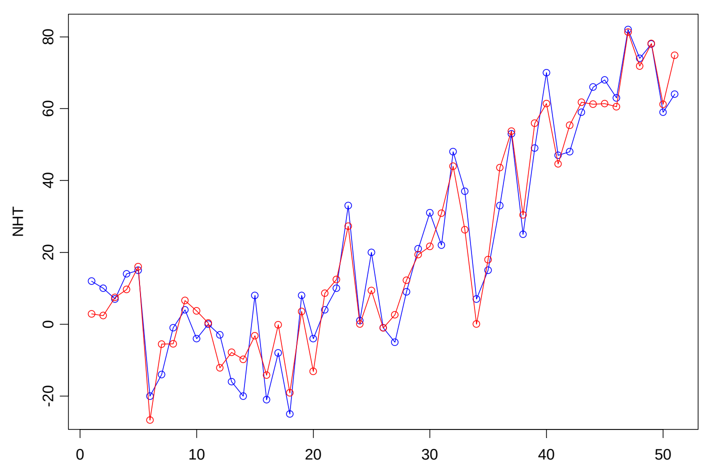
Y-W White Noise Covariance Matrix, V

.127399E+03	1.555684	-16.446952
1.555684	.110922	.357138
-16.446952	.357138	.131796E+03

Note: a) The coefficients of AR matrices are near zero in first column.
b) The residuals look good.

27

Observed data versus 1-step predicted values based on quadratic model plus VAR(2):
Data in blue, predictions in red.



28