MULTIVARIATE

ARMA PROCESSES

Stationary processes

A family X_t, $t\in\mathbb{Z}$, of random vectors $X_t:\Omega\to\mathbb{R}^k$ defined on a probability space $(\Omega,\mathcal{A},P)$ is called a **stationary** process if the mean vectors

$$EX_t=E\begin{pmatrix}X_{t1}\\ \vdots\\ X_{tk}\end{pmatrix}=\begin{pmatrix}EX_{t1}\\ \vdots\\ EX_{tk}\end{pmatrix}$$

and the autocovariance matrices

$$\mathrm{cov}(X_t,X_{t-h})=E(X_t-EX_t)(X_{t-h}-EX_{t-h})^T$$

are independent of t.

The **autocovariance function** of a weakly stationary process is defined by

$$\Gamma(h)=\mathrm{cov}(X_0,X_{0-h}).$$

Exercise: Show that $\Gamma(-h)=\Gamma^T(h)$. MS

A stationary process is called **white noise** if its auto-covariance function satisfies

$$\Gamma(h)=0 \quad \forall\ h\neq 0.$$

Any two components X_{tj} and $X_{(t+h)l}$ of white noise can be correlated with each other contemporaneously, although we have $\mathrm{cov}(X_{tj},X_{(t-h)l})=0$ for all $h\neq 0$.

Exercise: Show that each component of white noise is itself white noise. MW

First order autoregressive processes

A stationary process X_t, $t\in\mathbb{Z}$, is called a **first order autoregressive process** (or **AR(1) process**) if it can be expressed as

$$X_t=\Phi X_{t-1}+U_t,$$

where U_t, $t\in\mathbb{Z}$, is white noise with mean vector 0.

The AR(1) equation

$$\begin{pmatrix} X_{t1} \\ \vdots \\ X_{tk} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \cdots & \phi_{1k} \\ \vdots & \ddots & \vdots \\ \phi_{k1} & \cdots & \phi_{kk} \end{pmatrix} \begin{pmatrix} X_{(t-1)1} \\ \vdots \\ X_{(t-1)k} \end{pmatrix} + \begin{pmatrix} U_{t1} \\ \vdots \\ U_{tk} \end{pmatrix}$$

can also be written as

$$\begin{aligned} X_{t1}&=\phi_{11}X_{(t-1)1}+\ldots+\phi_{1k}X_{(t-1)k}+U_{t1}, \\ &\vdots \\ X_{tk}&=\phi_{k1}X_{(t-1)1}+\ldots+\phi_{kk}X_{(t-1)k}+U_{tk}. \end{aligned}$$

Thus, each component of an AR(1) process depends not only on lagged values of itself but also on lagged values of the other components.

Substituting $\Phi X_{t-2}+U_{t-1}$ for X_{t-1}, $\Phi X_{t-3}+U_3$ for X_{t-2}, … gives

$$
\begin{aligned}
X_t=\Phi X_{t-1}+U_t &= \Phi(\Phi X_{t-2}+U_{t-1})+U_t \\
&= \Phi^2 X_{t-2}+\Phi U_{t-1}+U_t \\
&= \Phi^2(\Phi X_{t-3}+U_{t-2})+\Phi U_{t-1}+U_t \\
&= \Phi^3 X_{t-3}+\Phi^2 U_{t-2}+\Phi U_{t-1}+U_t \\
&\;\;\vdots \\
&= \Phi^m X_{t-m}+\sum_{j=0}^{m-1}\Phi^j U_{t-j}
\end{aligned}
$$

Suppose that Φ is diagonalizable, i.e., there is an invertible matrix C such that $\Lambda=C^{-1}\Phi C$ is a diagonal matrix. It then follows from

$$\Lambda=C^{-1}\Phi C \Leftrightarrow \Phi=C\Lambda C^{-1}$$

that

$$
\begin{aligned}
&\Phi^2=\Phi\Phi=C\Lambda C^{-1}C\Lambda C^{-1}=C\Lambda^2C^{-1} \\
&\Phi^3=\Phi^2\Phi=C\Lambda^2C^{-1}C\Lambda C^{-1}=C\Lambda^3C^{-1} \\
&\quad\vdots \\
&\Phi^m=C\Lambda^mC^{-1}.
\end{aligned}
$$

Thus,

$$\Phi^m=C\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_k\end{pmatrix}^m C^{-1}=C\begin{pmatrix}\lambda_1^m & 0 & \cdots & 0\\ 0 & \lambda_2^m & & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_k^m\end{pmatrix}C^{-1}$$

will vanish as $m\to\infty$ only if the not necessarily real numbers $\lambda_1,\ldots,\lambda_k$ have modulus less than 1. MR

A $k \times k$ matrix Φ is diagonalizable if and only if it has k linearly independent eigenvectors (**diagonalization theorem**).

Proof:

Suppose that $c_1,\ldots,c_k$ are linearly independent eigen-vectors with eigenvalues $\lambda_1,\ldots,\lambda_k$. Then

$$\begin{aligned}\Phi &= \Phi(c_1,\ldots,c_k)(c_1,\ldots,c_k)^{-1}\\ &= (\Phi c_1,\ldots,\Phi c_k)(c_1,\ldots,c_k)^{-1}\\ &= (\lambda_1 c_1,\ldots,\lambda_k c_k)(c_1,\ldots,c_k)^{-1}\\ &= (c_1,\ldots,c_k)\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_k\end{pmatrix}(c_1,\ldots,c_k)^{-1}.\end{aligned}$$

If $\Phi = (c_1,\ldots,c_k)\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_k\end{pmatrix}(c_1,\ldots,c_k)^{-1}$, then

$$\begin{aligned}(\Phi c_1,\ldots,\Phi c_k) &= \Phi(c_1,\ldots,c_k)\\ &= (c_1,\ldots,c_k)\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_k\end{pmatrix}\\ &= (\lambda_1 c_1,\ldots,\lambda_k c_k).\end{aligned}$$

MD

Exercise: Show that if the 2×2 matrix Φ has 2 different eigenvalues, the corresponding eigenvectors will be linearly independent.

Solution: Since λ_1 and λ_2 are different, at least one of them, say λ_1, is not equal to zero. Assuming that the antithesis

$$c_1=\nu c_2 \text{ for some } \nu\neq 0$$

is valid, we obtain

$$\Phi c_1=\nu\Phi c_2,$$
$$\lambda_1 c_1=\nu\lambda_2 c_2,$$

and

$$c_1=\frac{\lambda_2}{\lambda_1}\nu c_2,$$

Since $\frac{\lambda_2}{\lambda_1}\neq 1$ and $c_1,c_2\neq 0$ this is in contradiction with the antithesis. ME

Exercise: Find the eigenvalues λ_1 and λ_2 of the matrix

$$\Phi=\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

MV

Hint: If c is an eigenvector with eigenvalue λ, i.e., $\Phi c=\lambda c$, then the homogeneous equation $(\lambda I-\Phi)c=0$ can only have a non-trivial solution $c\neq 0$, if the matrix $\lambda I-\Phi$ is non-invertible, hence λ must be a root of the characteristic polynomial $\det(\lambda I-\Phi)$.

Remark: The characteristic polynomial det($\lambda I-\Phi$) has degree k. According to the fundamental theorem of algebra it has therefore k (complex) roots, if each root is counted with its (algebraic) multiplicity. Since eigen-vectors corresponding to different eigenvalues are independent, Φ can only be non-diagonalizable if there exists an eigenvalue with algebraic multiplicity $m_a>1$ and geometric multiplicity $m_g<m_a$. The geometric multiplicity of an eigenvalue is the number of linearly independent eigenvectors with that eigenvalue.

Exercise: Show that the matrix

$$\Phi=\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

is non-diagonalizable. MN

The condition that all the eigenvalues of Φ are less than 1 in absolute value, i.e.,

$$|z| \geq 1 \quad \Rightarrow \det(\Phi\text{-}zI) \neq 0,$$

is equivalent to

$$|z| \geq 1 \quad \Rightarrow \det(-\tfrac{1}{z}(\Phi\text{-}zI)) \neq 0,$$

$$|z| \geq 1 \quad \Rightarrow \det(I-\tfrac{1}{z}\Phi) \neq 0,$$

and

$$|z| \leq 1 \quad \Rightarrow \det(I-z\Phi) \neq 0.$$

<u>Remark:</u> If all roots of the polynomial det(I-zΦ) lie outside of the unit circle, the sequence Φ, Φ^2, Φ^3, … is absolutely summable and

$$\sum_{j=0}^{\infty} \Phi^j U_{t-j}$$

converges (componentwise) in mean square to X_t.

Lag-operator notation

Using lag-operator notation the equation

$$X_t\text{-}\Phi X_{t-1}=U_t$$

can also be written as

$$(I\text{-}\Phi L)X_t=U_t,$$

where I-ΦL is a matrix-valued polynomial.

For example, in the bivariate case we have

$$I\text{-}\Phi L=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\text{-}\begin{pmatrix}\phi_{11} & \phi_{12}\\ \phi_{21} & \phi_{22}\end{pmatrix}L=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\text{-}\begin{pmatrix}\phi_{11}L & \phi_{12}L\\ \phi_{21}L & \phi_{22}L\end{pmatrix}$$

$$=\begin{pmatrix}1-\phi_{11}L & -\phi_{12}L\\ -\phi_{21}L & 1-\phi_{22}L\end{pmatrix}$$

and

$$(I\text{-}\Phi L)X_t=\begin{pmatrix}1-\phi_{11}L & -\phi_{12}L\\ -\phi_{21}L & 1-\phi_{22}L\end{pmatrix}\begin{pmatrix}X_{t1}\\ X_{t2}\end{pmatrix}$$

$$=\begin{pmatrix}(1-\phi_{11}L)X_{t1}-\phi_{12}LX_{t2}\\ -\phi_{21}LX_{t1}+(1-\phi_{22}L)X_{t2}\end{pmatrix}$$

$$=\begin{pmatrix}X_{t1}-\phi_{11}X_{(t-1)1}-\phi_{12}X_{(t-1)2}\\ X_{t2}-\phi_{21}X_{(t-1)1}-\phi_{22}X_{(t-1)2}\end{pmatrix}.$$

Autoregressive processes

A stationary process X_t, $t \in \mathbb{Z}$, is called an **autoregressive process of order p** (or **AR(p) process**) if it can be expressed as

$$X_t=\Phi_1 X_{t-1}+\ldots+\Phi_p X_{t-p}+U_t,$$

or, equivalently, as

$$X_t-\Phi_1 X_{t-1}-\ldots-\Phi_p X_{t-p}=U_t,$$

where U_t, $t \in \mathbb{Z}$, is white noise with mean vector 0.

Using lag-operator notation, the latter equation can also be written as

$$\Phi(L)X_t=U_t,$$

where

$$\Phi(L)=I-\Phi_1 L-\ldots-\Phi_p L^p$$

is a matrix-valued polynomial.

Yule-Walker equations

Postmultiplying the difference equation

$$X_t=\Phi_1 X_{t-1}+\ldots+\Phi_p X_{t-p}+U_t,$$

by X_{t-h}^T, h=0,1,2,…, and taking expectations we obtain the **Yule-Walker equations**:

$$
\begin{aligned}
\Gamma(0)&=\Phi_1\Gamma(-1)+\ldots+\Phi_p\Gamma(-p)+\Sigma,\\
\Gamma(1)&=\Phi_1\Gamma(0)+\ldots+\Phi_p\Gamma(-p+1),\\
&\vdots\\
\Gamma(p)&=\Phi_1\Gamma(p-1)+\ldots+\Phi_p\Gamma(0),\\
\Gamma(p+1)&=\Phi_1\Gamma(p)+\ldots+\Phi_p\Gamma(-1),\\
&\vdots
\end{aligned}
$$

When the parameter matrices $\Phi_1,\ldots,\Phi_p$, and Σ are given and the autocovariance matrices are unknown we can solve the first p+1 equations for $\Gamma(0),\ldots,\Gamma(p)$ and determine $\Gamma(p+1)$, $\Gamma(p+2),\ldots$ recursively.

Note that $\Gamma(-h)=\Gamma^T(h)$.

When the parameter matrices are unknown we may replace in the first p+1 equations $\Gamma(0),\ldots,\Gamma(p)$ by their sample counterparts $\hat{\Gamma}(0),\ldots,\hat{\Gamma}(p)$ and solve for $\Phi_1,\ldots,\Phi_p$, and Σ. The resulting estimators are called the **Yule-Walker estimators**.

Linear processes

Let $(X_t)_{t\in\mathbb{Z}}$ be a general **linear process** represented by

$$X_t = \sum_{j=-\infty}^{\infty} \Psi_j U_{t-j},$$

where $EU_t=0$, $\mathrm{var}(U_t)=\Sigma$, and $\mathrm{cov}(U_s,U_t)=0$ if $s\neq t$.

We have

$$EX_t = \sum_{j=-\infty}^{\infty} \Psi_j EU_{t-j} = 0$$

and

$$\begin{aligned}\mathrm{cov}(X_t,X_{t-k}) &= E\sum_{j=-\infty}^{\infty} \Psi_j U_{t-j}\left(\sum_{j=-\infty}^{\infty} \Psi_j U_{(t-k)-j}\right)^T \\ &= \sum_{r=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} \Psi_r EU_{t-r} U^T_{t-(j+k)} \Psi_j^T \\ &= \sum_{j=-\infty}^{\infty} \Psi_{j+k}\Sigma\Psi_j^T,\end{aligned}$$

respectively.

Since neither EX_t nor $\mathrm{cov}(X_t,X_{t-k})$ depend on t, the process $(X_t)_{t\in\mathbb{Z}}$ is weakly stationary. PL

The spectral density of a linear process

The spectral density $f_U(\omega)$ of white noise U_t, $t\in\mathbb{Z}$, with $\mathrm{var}(U_t)=\Sigma$ is given by

$$f_U(\omega)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{-i\omega k}\,\Gamma_U(k)=\frac{1}{2\pi}\Sigma$$

and the spectral density $f_X(\omega)$ of a linear process X_t, $t\in\mathbb{Z}$, represented by

$$X_t=\sum_{j=-\infty}^{\infty}\Psi_j U_{t-j}=\Big(\sum_{j=-\infty}^{\infty}\Psi_j L^j\Big)U_t$$

is given by

$$\begin{aligned}
f_X(\omega)&=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{-i\omega k}\,\Gamma_X(k)\\
&=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\Psi_{j+k}\Sigma\Psi_j^T e^{-i\omega k}\\
&=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\Psi_j\Sigma\Psi_{j-k}^T e^{-i\omega k}\\
&=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\Psi_j\Sigma\Psi_{j-k}^T e^{-i\omega j}e^{i\omega(j-k)}\\
&=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\left(e^{-i\omega j}\Psi_j\Sigma\right)\Psi_{j-k}^T e^{i\omega(j-k)}\\
&=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\Psi_j e^{-i\omega j}\,\Sigma\sum_{k=-\infty}^{\infty}\Psi_{j+k}^T e^{i\omega(j+k)}\\
&=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\Psi_j e^{-i\omega j}\,\Sigma\left(\sum_{k=-\infty}^{\infty}\Psi_k e^{-i\omega k}\right)^{*}.
\end{aligned}$$

PD

The spectral density of AR processes

The spectral density of an AR(1) process X_t, $t\in\mathbb{Z}$, with causal representation

$$(I-\Phi L)X_t=U_t$$

and MA(∞) representation

$$X_t=\left(\sum_{j=0}^{\infty}\Phi^j L^j\right)U_t$$

is given by

$$f_X(\omega)=\tfrac{1}{2\pi}\sum_{j=0}^{\infty}\Phi^j e^{-i\omega j}\,\Sigma\left(\sum_{k=0}^{\infty}\Phi^k e^{-i\omega k}\right)^*$$
$$=\tfrac{1}{2\pi}(I-\Phi e^{-i\omega})^{-1}\,\Sigma\left((I-\Phi e^{-i\omega})^{-1}\right)^*.$$

Exercise: Derive the sum formula

$$\sum_{j=0}^{n}\Phi^j=(I-\Phi^{n+1})(I-\Phi)^{-1}$$

for a geometric series of matrices. PG

Hint: Multiply each side of the equation by I-Φ.

Remark: Moreover, if all eigenvalues of Φ have modulus less than 1, we have

$$\sum_{j=0}^{\infty}\Phi^j=(I-\Phi)^{-1}.$$

Exercise: Show that

$$\sum_{j=0}^{\infty} \Phi^{j} e^{-i\omega j} = (I-\Phi e^{-i\omega})^{-1},$$

if all eigenvalues of Φ have modulus less than 1. PV

Analogously, the spectral density of an AR(p) process X_t, $t \in \mathbb{Z}$, represented by

$$(I-\Phi_1 L-\ldots-\Phi_p L^p)X_t=U_t$$

is given by

$$f_X(\omega)=\frac{1}{2\pi}\Phi^{-1}(e^{-i\omega})\,\Sigma\,(\Phi^{-1}(e^{-i\omega}))^{*},$$

where

$$\Phi^{-1}(e^{-i\omega})=(I-\Phi_1 e^{-i\omega}-\ldots-\Phi_p e^{-i\omega p})^{-1}.$$

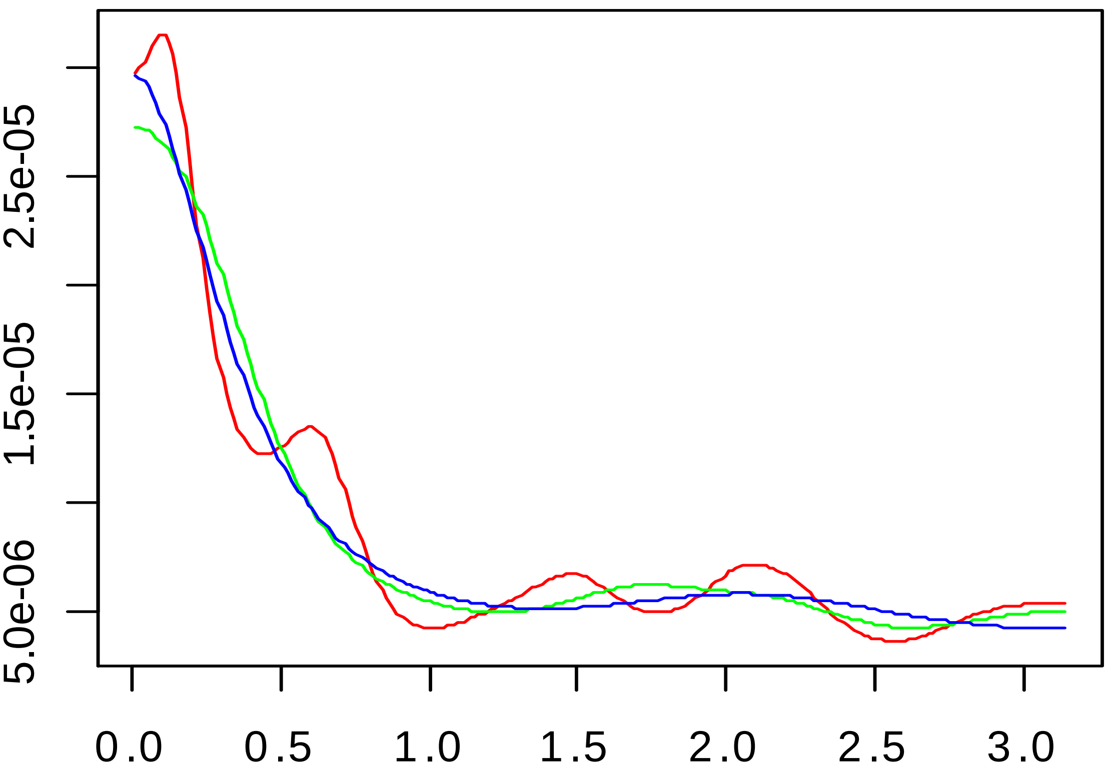
Unemployment and industrial production

To reexamine the relationship between changes in the industrial production (**dIP**) and changes in the duration of unemployment (**dUP**) with parametric methods we write a simple R function (**var.spev**) for the calculation of the spectral density of a vector autoregressive process.

```
var.spec <- function(fr,X.p)
# fr    … vector of frequencies
# X.p … AR(p) model estimated by R base function ar
  { p <- X.p$order; nf<-length(fr)
    im=complex(real=0,imaginary=1); pi2 <- 2*pi
    sigma <- X.p$var.pred; k <- length(sigma[1,])
    Id <- diag(1,nrow=k,ncol=k)
    sp <- array(dim=c(nf,k,k))
    for (w in 1:nf)
      { A <- Id
        for (l in 1:p) A <- A-X.p$ar[l,,]*exp(-im*fr[w]*l)
        A <- solve(A)
        sp[w,,] <- A%*%sigma%*%t(Conj(A))/pi2
      }
    return(sp)
  }
```

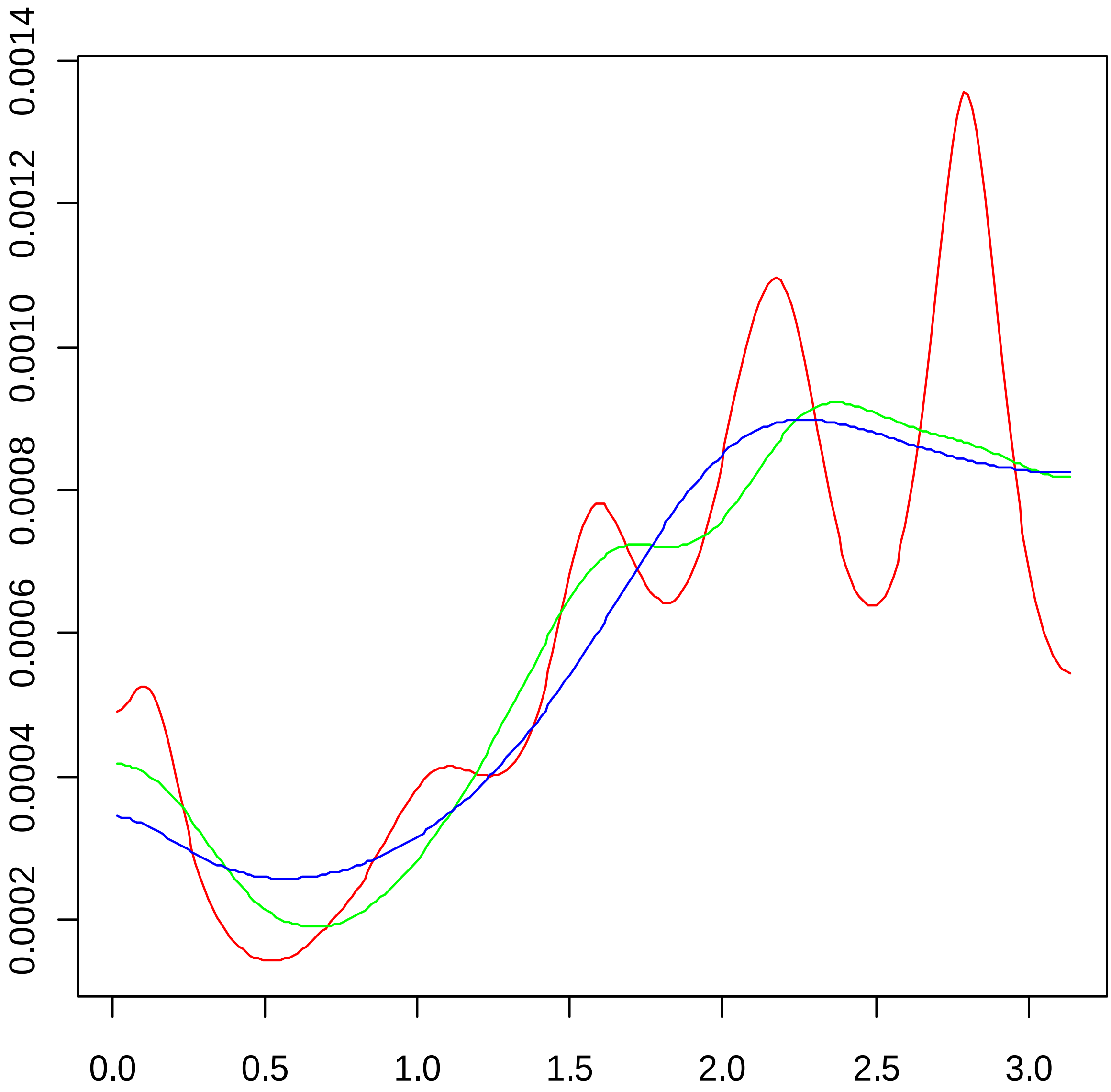
```
# Estimation of AR models of order p=3, 5, and 10
# aic=F … p is not selected automatically
# defaults: demean=T, method="yule-walker"
X.3 <- ar(X,order.max=3,aic=F)
X.5 <- ar(X,order.max=5,aic=F)
X.10 <- ar(X,order.max=10,aic=F)

X.3$var.pred # variance not explained by AR model
              dIP           dUE
dIP  4.283616e-05 -2.438674e-05
dUE -2.438674e-05  2.860293e-03

dim(X.3$ar) # dimension of array with AR coefficients
[1] 3 2 2

X.3$ar[1,,]; X.3$ar[2,,]; X.3$ar[3,,] # AR coefficients
           dIP          dUE
dIP  0.2823639 -0.003071477
dUE -1.7576082 -0.449250364
           dIP          dUE
dIP  0.1214053  0.001396195
dUE -1.7922100 -0.224859329
           dIP          dUE
dIP  0.1191343  0.002687604
dUE -0.5055041 -0.033355575
```

```
# Parametric spectral analysis
X.3.sp <- var.spec(fr,X.3)
X.5.sp <- var.spec(fr,X.5)
X.10.sp <- var.spec(fr,X.10)

# Spectrum of dIP
plot(fr,X.10.sp[,1,1],type="l",col="red",
      xlab=" ",ylab=" ")
lines(fr,X.5.sp[,1,1],col="green")
lines(fr,X.3.sp[,1,1],col="blue")
```

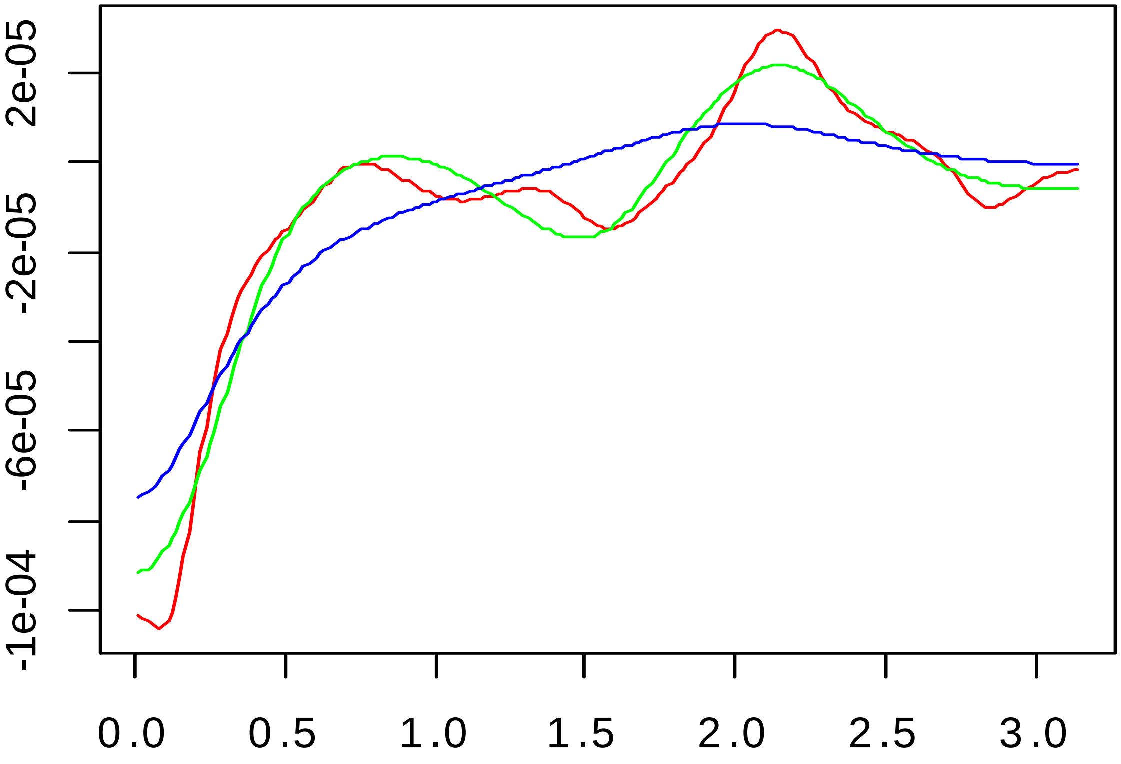


Spectrum of dUE

```
plot(fr,X.10.sp[,2,2],type="l",col="red",
      xlab=" ",ylab=" ")
lines(fr,X.5.sp[,2,2],col="green")
lines(fr,X.3.sp[,2,2],col="blue")
```

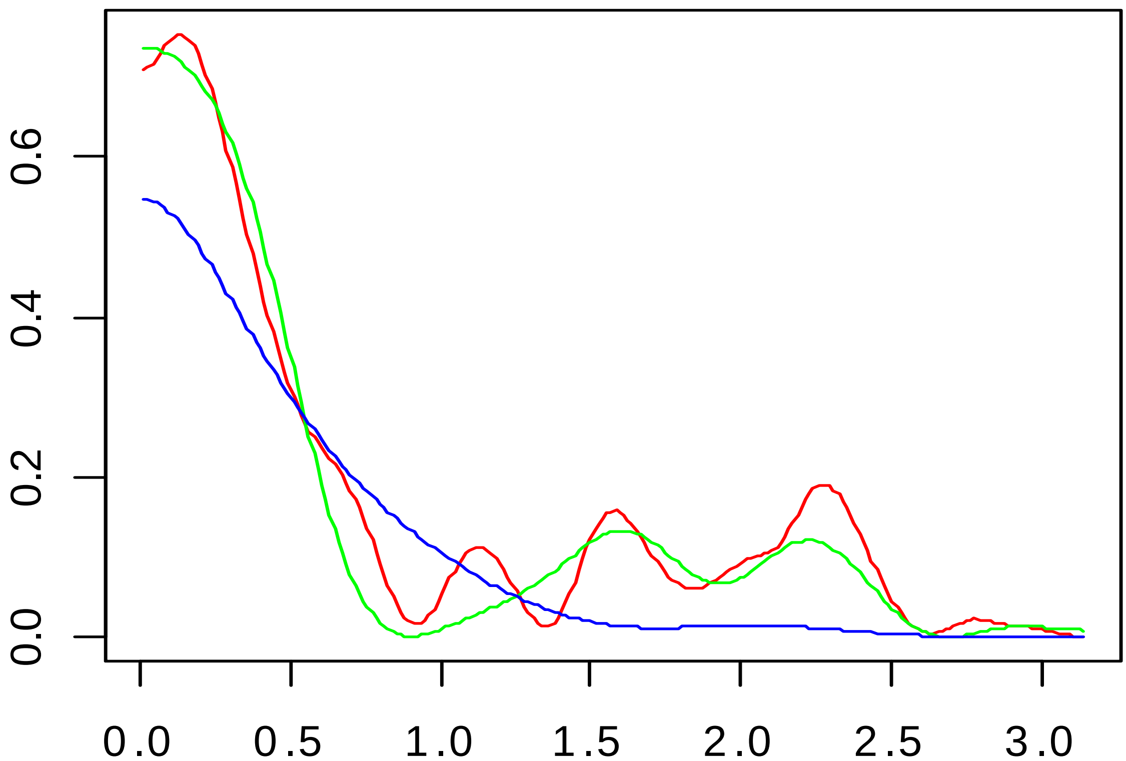


```
# Cospectrum
plot(fr,Re(X.10.sp[,1,2]),type="l",col="red",
    xlab=" ",ylab=" ")
lines(fr, Re(X.5.sp[,1,2]),col="green")
lines(fr, Re(X.3.sp[,1,2]),col="blue")
```



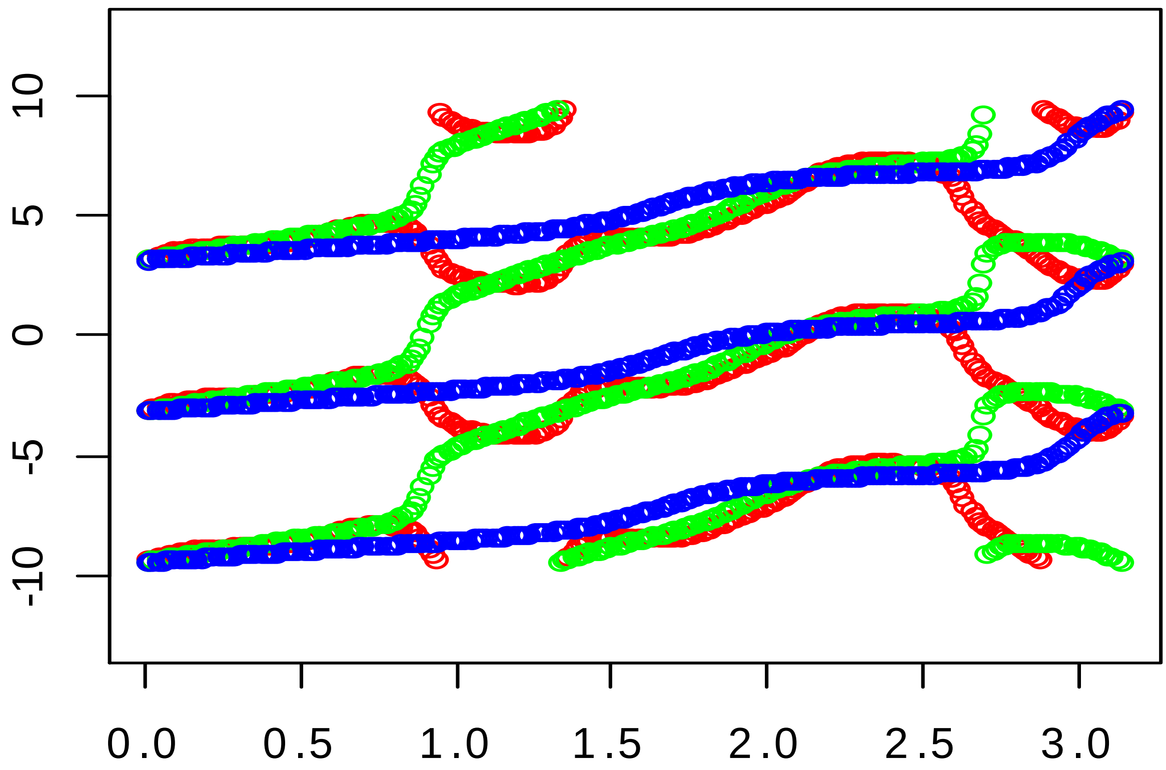
The overall negative relationship between the two variables is mainly due to the low frequencies.

```
# Squared coherency
plot(fr,Mod(X.10.sp[,1,2])^2
       /( X.10.sp[,1,1]*X.10.sp[,2,2]),
       type="l",col="red",xlab=" ",ylab=" ")
lines(fr,Mod(X.5.sp[,1,2])^2
      /( X.5.sp[,1,1]*X.5.sp[,2,2]),col="green")
lines(fr,Mod(X.3.sp[,1,2])^2
      /( X.3.sp[,1,1]*X.3.sp[,2,2]),col="blue")
```



The linear relationship between the two variables is strongest at the low frequencies.

```
# Phase spectrum
plot(fr,Arg(X.10.sp[,1,2]),ylim=c(-4*pi,4*pi),
     col="red",xlab=" ",ylab=" ")
lines(fr,Arg(X.10.sp[,1,2])-2*pi,"p",col="red")
lines(fr,Arg(X.10.sp[,1,2])+2*pi,"p",col="red")
lines(fr,Arg(X.5.sp[,1,2]),"p",col="green")
lines(fr,Arg(X.5.sp[,1,2])-2*pi,"p",col="green")
lines(fr,Arg(X.5.sp[,1,2])+2*pi,"p",col="green")
lines(fr,Arg(X.3.sp[,1,2]),"p",col="blue")
lines(fr,Arg(X.3.sp[,1,2])-2*pi,"p",col="blue")
lines(fr,Arg(X.3.sp[,1,2])+2*pi,"p",col="blue")
```



At the low frequencies, there is a slight positive slope indicating that the first series (**dIP**) is leading.

ARMA processes

A stationary process X_t, $t\in\mathbb{Z}$, is called an **autoregressive moving average process of order (p,q)** (or **ARMA(p,q) process**) if it can be expressed as

$$X_t=\Phi_1X_{t-1}+\ldots+\Phi_pX_{t-p}+U_t+\Theta_1U_{t-1}+\ldots+\Theta_qU_{t-q}$$

or, equivalently, as

$$X_t-\Phi_1X_{t-1}-\ldots-\Phi_pX_{t-p}=U_t+\Theta_1U_{t-1}+\ldots+\Theta_qU_{t-q},$$

where U_t, $t\in\mathbb{Z}$, is white noise with mean vector 0.

Using lag-operator notation, the latter equation can also be written as

$$\Phi(L)X_t=\Theta(L)U_t,$$

where

$$\Phi(L)=I-\Phi_1L-\ldots-\Phi_pL^p$$

and

$$\Theta(L)=I+\Theta_1L+\ldots+\Theta_qL^q$$

are matrix-valued polynomials.

An ARMA(p,0) process is an AR(p) process. An ARMA(0,q) process is also called a **moving average process of order q** (or **MA(q) process**).

Causality and invertibility

The ARMA(p,q) equation

$$(I-\Phi_1 L-\ldots-\Phi_p L^p)X_t=(I+\Theta_1 L+\ldots+\Theta_q L^q)U_t$$

is said to be **causal** if

$$|z|\le 1 \Rightarrow \det(I-z\Phi_1-\ldots-z^p\Phi_p)\ne 0.$$

It is said to be **invertible** if

$$|z|\le 1 \Rightarrow \det(I+z\Theta+\ldots+z^q\Theta_q)\ne 0.$$

Exercise: Show that the bivariate AR(1) process

$$\begin{pmatrix} X_{t1} \\ X_{t2} \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_{(t-1)1} \\ X_{(t-1)2} \end{pmatrix} = \begin{pmatrix} U_{t1} \\ U_{t2} \end{pmatrix}$$

is causal and invertible. PC

Exercise: Show that the bivariate MA(2) process

$$\begin{pmatrix} X_{t1} \\ X_{t2} \end{pmatrix} = \begin{pmatrix} U_{t1} \\ U_{t2} \end{pmatrix} + \begin{pmatrix} 0 & -\frac{1}{2} \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} U_{(t-1)1} \\ U_{(t-1)2} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} U_{(t-2)1} \\ U_{(t-2)2} \end{pmatrix}$$

is causal and invertible. PI

Exercise: Show that the bivariate ARMA(1,1) process

$$\begin{pmatrix} X_{t1} \\ X_{t2} \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 2 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} X_{(t-1)1} \\ X_{(t-1)2} \end{pmatrix} = \begin{pmatrix} U_{t1} \\ U_{t2} \end{pmatrix} + \begin{pmatrix} 0 & \frac{5}{3} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_{(t-1)1} \\ U_{(t-1)2} \end{pmatrix}$$

is causal and invertible. PA

Identifiability

It does not make sense to estimate the parameter matrices $\Phi_1,\ldots,\Phi_p,\Theta_1,\ldots,\Theta_q$, and Σ of an ARMA(p,q) process if they are not unique.

To ensure identifiability in the univariate case, where $\Phi(L)$ and $\Theta(L)$ are just scalar polynomials, we must require, in addition to causality and invertibility, that $\Phi(z)$ and $\Theta(z)$ have no common zeros. For example, the equation

$$(1-\tfrac{1}{4}L^2)X_t=(1+\tfrac{1}{2}L)U_t$$

can be written more parsimoniously as

$$(1-\tfrac{1}{2}L)X_t=U_t,$$

because the polynomials $1-\frac{1}{4}z^2=(1+\frac{1}{2}z)(1-\frac{1}{2}z)$ and $1+\frac{1}{2}z$ have a common zero.

In the multivariate case, the matrix-valued polynomials $\Phi(z)$ and $\Theta(z)$ can have a common left factor even if $\det(\Phi(z))$ and $\det(\Theta(z))$ have no common zero. To avoid the difficulties involved in the identification of multivariate ARMA processes, many time series analysts use only multivariate AR models for the modeling of multivariate time series.

Exercise: Show that the equation

$$\begin{pmatrix} X_{t1} \\ X_{t2} \end{pmatrix} - \begin{pmatrix} 0 & \phi+\theta \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_{(t-1)1} \\ X_{(t-1)2} \end{pmatrix} = \begin{pmatrix} U_{t1} \\ U_{t2} \end{pmatrix} + \begin{pmatrix} 0 & -\theta \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_{(t-1)1} \\ U_{(t-1)2} \end{pmatrix}$$

can be written more parsimoniously as

$$\begin{pmatrix} X_{t1} \\ X_{t2} \end{pmatrix} - \begin{pmatrix} 0 & \phi \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_{(t-1)1} \\ X_{(t-1)2} \end{pmatrix} = \begin{pmatrix} U_{t1} \\ U_{t2} \end{pmatrix}$$

although the polynomials

$$\det(\Phi(z)) = \det(I - \begin{pmatrix} 0 & \phi+\theta \\ 0 & 0 \end{pmatrix} z)$$

and

$$\det(\Theta(z)) = \det(I + \begin{pmatrix} 0 & -\theta \\ 0 & 0 \end{pmatrix} z)$$

have no common zero. PU

Hint: Multiply both $\Phi(z)$ and $\Theta(z)$ by $\Theta^{-1}(z) = \begin{pmatrix} 1 & \theta z \\ 0 & 1 \end{pmatrix}$.

Remark: The inverse of the matrix-valued polynomial $\Theta(z)$ is also a matrix-valued polynomial. Its determinant is a constant unequal to zero. Such a matrix-valued polynomial is called **unimodular**.