MULTIVARIATE TIME SERIES ANALYSIS AND FORECASTING

Manfred Deistler

Econometrics and Systems Theory

Institute for Mathematical Methods in Economics
University of Technology Vienna
Singapore, May 2004

Introduction

Time Series: Observations ordered in time, information contained in ordering: Results are not permutation-in-variant in general.

Discrete time, equally spaced data: $x_t, t = 1 \ldots T; x_t \in \mathbf{R}^n$

Main questions in TSA:

- ✓ Data driven modelling (system identification)
- ✓ Signal and feature extraction (e.g. seasonal adjustment)

Theory and methods concern:

- ✓ Model classes
- ✓ Estimation and inference
- ✓ Model selection
- ✓ Model evaluation

Areas of application:

- ✓ Signal processing e.g. speech, sonar and radar signals
- ✓ Data driven modelling for simulation and control of technical systems and processes; monitoring
- ✓ Time series econometrics: Macroeconometrics, finance econometrics, applications to marketing and firm data
- ✓ Medicine and biology: Genetics, EEG data, monitoring ...

Tutorial Lecture 1: Stationary Processes

Stationary processes are an important model class for time series

<u>Def:</u> A stochastic process $(x_t)(t \in \mathbb{Z})$, $x_t : \Omega \to \mathbb{R}^n$, is called (wide sense stationary if

(i) $E\, x_t' x_t < \infty$ for all t

(ii) $E\, x_t = m = const$ for all t

(iii) $\gamma(s) = E(x_{t+s} - m)(x_t - m)'$ does not depend on t

“Shift invariance” of first and second moments

$\gamma : \mathbf{Z} \to \mathbf{R}^{nxn}$ covariance function of (x_t)

contains all linear dependence relations between process variables

$$\Gamma_T = \begin{pmatrix} \gamma(0) & \dots & \gamma(-T+1) \\ \dots & \dots & \dots \\ \gamma(T-1) & \dots & \gamma(0) \end{pmatrix}$$

A function $\gamma : \mathbf{Z} \to \mathbb{R}^{nxn}$ is called <u>nonnegative definite</u> if $\Gamma_T \geq 0 \quad \forall T$

<u>Mathematical characterization of covariance functions</u> of stationary processes:
γ is a covariance function if and only if γ is nonnegative definite

Examples for stationary processes

(1) White noise

$E\varepsilon_t = 0 \qquad E\varepsilon_s\varepsilon_t' = \delta_{st}\,\Sigma, \quad \Sigma \geq 0$ no (linear) memory

(2) Moving average (MA) process

$x_t = \sum_{j=0}^{q} b_j \varepsilon_{t-j}; \qquad b_j \in \mathbb{R}^{nxm}$ finite memory

(3) Infinite moving average process

$x_t = \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$ large class of stationary processes

(4) (Stationary) Autoregressive (AR) process

Steady state solution of stable VDE of the form

$$\sum_{j=0}^{p} a_j x_{t-j} = \varepsilon_t; \quad a_j \in \square^{nxn} \quad \det a(z) \neq 0 \quad |z| \leq 1, a(z) = \sum_{j=0}^{p} a_j z^j$$

(5) (Stationary) ARMA process

Steady state solution of stable VDE of the form

$$a(z)x_t = b(z)\varepsilon_t \;\; z \ldots backwardshift; \quad b(z) = \sum_{j\infty}^{q} b_j z^j$$

(6) <u>Harmonic process</u>

$$x_t = \sum_{j=0}^{h} e^{i\lambda_j t} z_j = \sum_{j=1}^{\lfloor (h+1)/2 \rfloor} a_j \cos \Upsilon_j t + b_j \sin \Upsilon_j t$$

where $\lambda_j \in [-\pi, \pi]$ (angular) frequencies,

$$\Upsilon_j = \lambda_{j+\lfloor h/2 \rfloor} j = 1 \ldots \lfloor (h+1)/2 \rfloor \in [0, T]$$

$$z_j : \Omega \to \mathbf{C}^n, a_j, b_j : \Omega \to \square^n$$

π : Nyquist frequency; (x_t) is a weighted sum of harmonic oscillations with random weights (amplitudes and phases)

<u>Stationarity conditions:</u>

$$Ez_j^* z_j < \infty, Ez_j = \begin{cases} 0 & for \quad j : \lambda_j \neq 0 \\ Ex_t & for \quad j : \lambda_j = 0 \end{cases}, Ez_j z_1^* = 0 \quad j \neq 1$$

Spectral distribution function for a harmonic process

$$F: \quad [-\pi, \pi] \rightarrow \square^{nxn} : F(\lambda) = \sum_{j:\lambda_j \leq \lambda} F_j \quad ; F_j = E z_j z_j^*$$

$$F \leftrightarrow \gamma$$

has the same information as γ about the process, however displayed in a different way

Spectral representation of stationary process

Every stationary process is the (pointwise in t) limit of a sequence of harmonic processes:

<u>Theorem:</u> Every stationary process (x_t) can be represented as

$$x_t = \int_{[-\pi,\pi]} e^{i\lambda t} dz(\lambda)$$

where $(z(\lambda) \mid \lambda \in [-\pi,\pi])$, $\quad z(\lambda) : \Omega \to \mathbb{C}^n$ satisfies

$$z(-\pi) = 0, \quad z(\pi) = x_0, \quad Ez(\lambda)z^*(\lambda) < \infty, \quad \lim_{\varepsilon \downarrow 0} z(\lambda + \varepsilon) = z(\lambda),$$

$$E\left\{(z(\lambda_4) - z(\lambda_3))(z(\lambda_2) - z(\lambda_1))^*\right\} = 0 \quad \text{for } \lambda_1 < \lambda_2 \le \lambda_3 < \lambda_4$$

and is unique for given (x_t)

Spectral distribution of a general stationary process

$$F : [-\pi, \pi] \to \square^{nxn} \quad : F(\lambda) = Ez(\lambda)z^{*}(\lambda)$$

Spectral density

If $\sum \| \gamma(t) \| < \infty$

then there exists a function $f : [-\pi, \pi] \to \square^{nxn}$

s.t. $F(\lambda) = \int_{-\pi}^{\lambda} f(\omega)$ dw, called the <u>spectral density</u>

We have

$$\gamma(t) = \int e^{i\lambda t} f(\lambda) d\lambda$$

$$f(\lambda) = (2\pi)^{-1} . \sum_{t=-\infty}^{\infty} e^{-i\lambda t} \gamma(t)$$

f is characterized by $f \geq 0 \quad \lambda$ a.e., $|| \int f(\lambda) d\lambda || < \infty$ and $f(\lambda) = f(-\lambda)'$

In particular we have $\gamma(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda$: Variance decomposition

The diagonal elements of f show the contributions of the frequency bands to the variance of the respective component process and the off-diagonal elements show the frequency band specific covariances and expected phase shifts between component processes.

Parametric Estimation	Seminonparametric Estimation	Nonparametric Estimation
e.g. *AR* estimation for given p	e.g. *AR* estimator where in addition p is estimated	e.g. Windowed spectral estimation

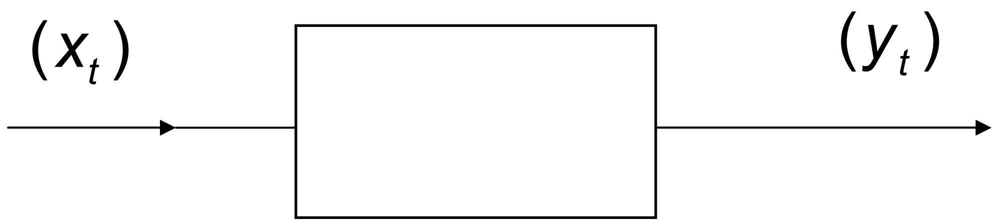
"The curse of dimensionality": E.g. for *AR* estimation (with given p) the dimension of the parameter space is $n^2 p$ (for the a_j) plus

$$\frac{n(n+1)}{2} \text{ for } \Sigma$$

Linear transformations of stationary processes

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j} \quad ; a_j \in \square\, nxm, \overset{\mathbf{n}}{\sum} \| aj \|^2 < \infty, \quad (x_t) \text{ stationary}$$

linear, dynamic, time invariant, stable system



Weighting function $(a_j \mid j \in \mathbf{Z})$

Causality: $a_j = 0 \quad j < 0$

$$y_t = \int e^{i\lambda t} dz_y(\lambda) = \sum_{j=-\infty}^{\infty} a_j \int e^{i\lambda(t-j)} dz_x(\lambda) = \int e^{i\lambda t} \underbrace{\left(\sum_{j_0-\infty}^{\infty} a_j e^{-i\lambda j} \right)}_{\text{transfer function}} dz_x(\lambda)$$

frequency-specific gain and phase shift

$$k(z) = \sum_{j=-\infty}^{\infty} a_j z^j \leftrightarrow (a_j \mid j \in \mathbf{Z})$$

Transformation of second moments

$$f_{yx}(\lambda) = k(e^{-i\lambda}) f_x(\lambda)$$

$$f_y(\lambda) = k(e^{-i\lambda}) f_x(\lambda) k^*(e^{-i\lambda})$$

Solution of linear vector difference equations (VDE's)

$$a_o y_t + a_1 y_{t-1} + \ldots + a_p y_{t-p} = b_0 x_t + \ldots + b_q x_{t-q} \;;\; a_j \in \mathbb{R}^{nxn}$$

$$b_j \in \mathbb{R}^{nxm}, t \in \mathbb{Z}$$

or:

$$a(z) y_t = b(z) x_t$$

where $a(z) = \sum_{j=0}^{p} a_j z^j \;, \quad b(z) = \sum_{j=0}^{q} b_j z^j$

$z : z \in \mathbb{C}$ as well as backward-shift: $z(y_t \mid t \in \mathbb{Z}) = (y_{t-1} \mid t \in \mathbb{Z})$

Steady state solution: $z - transform$

If $\det a(z) \neq 0 \quad |z| \leq 1$ then there exists a causal stable solution

$$y_t = \sum_{j=0}^{\infty} k_j x_{t-j}$$

where $\sum_{0}^{\infty} k_j z^j = k(z) = a^{-1}(z) b(z) = (\det a(z))^{-1} adj(a(z)) b(z)\,;\; z \in \square$

Forecasting for stationary processes

Problem: Approximation of a future value $x_{t+h}, h > 0$ from the past $x_s, s \leq t$

Linear least squares forecasting:

$$\min_{a_j \in \mathbb{R}^{nxn}} E(x_{t+h} - \sum_{j\geq 0} a_j x_{t-j})^* (x_{t+h} - \sum a_j x_{t-j})$$

Projection interpretation

Prediction from a finite past; $x_t, x_{t-1}, \ldots, x_{t-r}$

$$E(x_{t+h} - \sum_{j=0}^{r} a_j x_{t-j}) x_{t-s}' = 0 \text{ , } s = 0, \ldots, r$$

leads to

$$(a_o,\ldots,a_r)\begin{pmatrix}\gamma(0) & \cdots & \gamma(r)\\ & \ddots & \\ \gamma(-r) & & \gamma(0)\end{pmatrix}=(\gamma(h),\ldots,\gamma(h+r))$$

$$\hat{x}_{t,h}=\sum a_j x_{t-j} \qquad \text{Predictor}$$

<u>Prediction from an infinite past</u>; $x_t, x_{t-1},\ldots$

A stationary process is called (linearly) <u>singular</u> if

$$\hat{x}_{t,h}=x_{t+h} \qquad \text{for some and hence for all } t,h>0$$

Here $\hat{x}_{t,h}$ denotes the best linear least squares predictor from the infinite past.

A stationary process is called (linearly) regular if

$$\lim_{h\to\infty} \hat{x}_{t,h} = 0$$

for one and hence for all t.

Theorem (Wold)

(i) Every stationary process (x_t) can be represented in a unique way as $x_t = y_t + z_t$ where (y_t) is regular, (z_t) is singular, $Ey_t z_s' = 0$ and y_t and z_t are causal linear transformations of (x_t)

(ii) Every regular process (y_t) can be represented as:

$y_t = \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}$; $\sum \| k_j \|^2 < \infty$, (ε_t) white noise and where ε_t is a causal linear transformation of (y_t)

Consequences for forecasting:

(i) (y_t) and (z_t) can be forecasted separately

(ii) for the regular process (y_t) we have:

$$y_{t+h} = \sum_{j=0}^{\infty} k_j \varepsilon_{t+h-j} = \underbrace{\sum_{j=h}^{\infty} k_j \varepsilon_{t+h-j}}_{predictor\ \hat{y}_{t,h}} + \underbrace{\sum_{j=0}^{h-1} k_j \varepsilon_{t+h-j}}_{prediction\ error}$$

Note: Every regular process can be forecasted with arbitray accuracy by an $(AR)MA$ process

How do we obtain the Wold representation: Spectral factorization

$$f_y = (2\pi)^{-1} k(e^{-i\lambda}) \Sigma k(e^{-i\lambda})^*, \ \Sigma = E\varepsilon_t \varepsilon_t'$$