TSA by JG

[TOC]

Intro

Definition A time series model for *the observed data* $\{x_t\}$ is a specification of the joint distributions of a sequnce of random variables $\{X_t\}$ of which $\{x_t\}$ is postulated to be a realization.

Definition (IID noise) A process $\{X_t\ ,\ t\ \in\ \mathbb{Z}\}$ is said to be a IID noise with mean 0 and variance σ^2, written

$$\{X_t\} \sim \mathrm{IID}(0, \sigma^2)$$

if the random variables X_t are independent and indentically distributed with $EX_t = 0$ and $\mathrm{Var}(X_t) = \sigma^2$

(IID = Independent and Identically Distributed) or $X_t \sim \text{i.i.d.}$ for short

obviously the binary process is $\mathrm{IID}(0{,}1)$ noise.

Definition Let $\{X_t\,,\, t\,\in\, T\}$ or $\{X_t\}$ for the laziness's sake, with $\mathrm{Var}(X_t) < \infty\, \forall t$
The mean function of $\{X_t\}$ is:

$$\mu_X(t) := E(X_t), \quad t\,\in\, T$$

The covariance funtion of X_t is:

$$\gamma_X(r, s) := \mathrm{Cov}(X_r, X_s)$$

Stationarity

Loosely speaking, a stochastic process is stationary, if its statistical properties do not change with time.

Definition(Stationary) The time series $\{X_t\,,\, t\,\in\, \mathbb{Z}\}$ is said to be (weakly) stationary if

- $\mathrm{Var}(X_t) < \infty$
- $\mu_X(t) = \mu$
- $\gamma_X(r, s) = \gamma_X(r + t, s + t)$

the last condition implies that $\gamma_X(r, s)$ is a funtion of $r - s$ then:

$$\gamma_X(h) := \gamma_X(h, 0)$$

The value "h" is referred to as the "lag"

Definition (ACVF) Let $\{X_t\}$ be a stationary time series. The autocovariance function (ACVF) of $\{X_t\}$ is

$$\gamma_X(h) = \mathrm{Cov}(X_{t+h}, X_t)$$

The autocorrelation function is

$$\rho_X(h) := \frac{\gamma_X(h)}{\gamma_X(0)}$$

Definition (White Noise) A process $\{X_t\}$ is said to be a white noise with mean μ and covariance σ^2 , written:

$$X_t \sim \mathrm{WN}(\mu, \sigma^2)$$

if $EX_t = \mu$ and $\gamma(h) = \begin{cases} \sigma^2 & h = 0 \\ 0 & h \neq 0 \end{cases}$

Trends & Seasonal Components

$$X_t = m_t + s_t + Y_t$$

be the "classical decomposition" model where:

- m_t is a slowly changing function (the "Trend Componet")
- s_t is a function with known (or given) period d (the "Seasonal Complement")
- Y_t is a stationary time series

> *Our aim is to estimate and extract the deterministic components m_t and s_t in hope that the residual component Y_t will turn out to be a stationary time series -- Jan Grandell*

No Seasonal Components

Assume that

$$X_t = m_t + Y_t, \quad t = 1, \cdots, n$$

where, without loss of generality, $EY_t = 0$.(That is beacause is $EY_t = \mu$ (the Y_t is stationary) then $X_t = (m_t + \mu) + (EY_t - \mu)$ is the form we mentioned)

Method 1 (Least Squares Estimation of m_t)
Method 2 (Smooting by means of a moving average)
Method 3 (Differencing to generate stationarity)

Trend and Seasonality

Let us go back to

$$X_t = m_t + s_t + Y_t$$

where $EY_t = 0$, $s_{t+d} = s_t$ and $\sum_{k=1}^{d} s_k = 0$. Assume that n/d is an integer.

Sometimes it is convenient to index the data by period and time-unit

$$x_{j,k} = x_{k+d(j-1)}, \quad k = 1, \cdots, d, \; j = 1, \cdots, \frac{n}{d}$$

Autocovariance

Strict stationarity

spectral density

Time Series Model TSM

Estimation

Estimation of μ

Estimation of $\gamma(\cdot)$ and $\rho(\cdot)$

Prediction

inference

Prediction of Random Variables PRV

Further Prediction

Further PRV

Predition for stationary time series PSTS

wold decomposition

Partial Correlation

PA Partial Autocorrelation

ARMA Processes

ACVF and how to cal

Prediction of ARMA

Spectral Analysis

Spectral Distribution

Spectral Representation of a time series

Predition in the frequency domain PiFD

interpolation and detection

The Itˆo intergral

Estimation of the spectral density

periodogram

Smoothing the Periodogram SP

Linear Filters

ARMA Processes

Estimation for ARMA models

Yule-Walker estimation

Burg's algorithm

innovation algorithm

hannan-Rissanen algorithm

Maximum Likelihood and Lead Square Estimation

Order Selection

unit roots

Multivariate Time Series MTS

Financial Time Series

Kalman Filtering

State-Space Representation

Prediction of Multivariate random variables

Appendix

A.1 Stochastic Processes

Definition (Stochastic Process) A stochastic process is a *family of random variables* $\{X_t\,,\,t\in T\}$ *defined on a probability space* $(\Omega,\mathcal{F},P)$ as follows:

- Ω is a set
- $\mathcal{F}$ is a σ-field i.e.
  - (a) $\emptyset\,\in\,\mathcal{F}$
  - (b) $A_i\,\in\,\mathcal{F}\quad\forall\, i\,\in\, I$ then $\bigcup_i A_i\,\in\,\mathcal{F}$
  - (c) $A\,\in\,\mathcal{F}$ then $A^c\,\in\,\mathcal{F}$
- P is a function $\mathcal{F}\to[0,1]$ satisfying:
  - $P(\Omega)=1$
  - $A_i\,\in\,\mathcal{F}\quad\forall\, i\,\in\, I$ and $A_i\cap A_j=\emptyset\quad\forall\, i\,\in\, I$ then $P(\bigcup_i A_i)=\sum_i P(A_i)$

proposition $P(A)+P(A^c)=1$

There are definitions on Cholton's Style:

(X is *random variable*) iff ((X is a function $\Omega\to\mathbb{R}$) and ($\{\omega\,\in\,\Omega: X(\omega)\le x\}\,\in\,\mathcal{F}$ for all $x\,\in\,\mathbb{R}$))

and we denoted $P(X\le x):=P(X^{-1}([-\infty,x])$ and $P(X<x):=P(X^{-1}([-\infty,x)$

($\{X_t\,,\,t\in T\}$ is *stochastic process*) iff (X_t is *random variable* for all $t\,\in\,T$)

and T is called *index* or *parameter set*

($\{X_t\,,\,t\in T\}$ is a *time series*) iff ($T\subset\mathbb{Z}$)

Definition (sample-path) The functions $\{X_t(\omega)\,,\,\omega\,\in\,\Omega\}$ on T are called realizations or *sample-path* of the process $\{X_t\,,\,t\,\in\,T\}$

$F_X, x\to P(X\le x)$ is called the distribution function of a random variable X
$F_{(\cdot)}:(\Omega\to\mathbb{R})\to(\mathbb{R}\to[0,1])$ is called the distribution

The case in higher demension is similar.

$X=(X_1,\,\cdots,X_n)^\top$ is a n-dim *random variable*, X_i is a *random variable* for $1\,\le\, i\,\le\, n$

Definition (The distribution of a stochastic process) let
$\mathcal{T}:=\{t\in T^n\,:\,t_i<t_j\}$
The (finitie-dimensional) distribution function are the family $\{F_t(\cdot)\,,\,t\,\in\,\mathcal{T}\}$,
$F_t(x)=P(X_{t_1}\le x_1,\cdots,X_{t_n}\le x_n)\quad t\,\in\,T^n\,,\,x\in\mathbb{R}$
The distribution of $\{X_t\,,\,t\in T\}$ is the family $\{F_t(\cdot)\,,\,t\,\in\,\mathcal{T}\}$

obviously $\mathcal{T}$ is a simplex
In a way, we can say:"$t\in T$ is a *series of time*". So the F_t is a distribution of a *series of time* of random variable i.e. the distribution of *stochastic process*. In conviencity, $F_t\sim\mathcal{T}\sim X_t\sim n$, where the symbol $\sim$ readed as 'related to'.

Theorem (Kolmogorov's existence theorem) The family $F_t(\cdot)\,,\,t\,\in\,\mathcal{T}$ are the distribution functions of some stochastic process iff for any n, $t=(t_1,\cdots,t_n)\,\in\,\mathcal{T}$, $x\,\in\,\mathbb{R}^n$ and $1\le k\le n$

$$\lim_{x_k\to\infty}F_t(x)=F_{t(k)}(x(k))$$

where $t(k)=(t_1,\cdots,t_{k-1},t_{k+1},\cdots,t_n)$ and $x(k)=(x_1,\cdots,x_{k-1},x_{k+1},\cdots,x_n)$ (It is to say, $t(k)$ is t *deleted the k-th variable* and $x(k)$ is x *deleted the k-th variable*)

$$\phi_t(u) = \int_{\mathbb{R}^n} e^{iu'x} F_t(dx_1, \cdots, dx_n)$$

be the characteristic function of F. Then KET can be restated as follow:

$$\lim_{u_i \to 0} \phi_t(u) = \phi_{t(i)}(u(i))$$

where $u(i)$ ana $t(i)$

(a n-dim r.v. Y is *normally distributed*) iff (($Y = AX + b$) where $A \in M_n(\mathbb{R})$, $X \sim N_n(0, 1)$ and $b \in \mathbb{R}^n$)

and apparently

$$\mu_Y := E(Y) := (E(Y_1), \cdots, E(Y_n))^\top = E(AX + b) = E(AX) + E(b) = A \cdot E(X) + b = A \cdot 0 + b = b$$

$$\Sigma_{YY} := \mathrm{Cov}(Y, Y) := E([Y - E(Y)][Y - E(Y)]') = E([AX + b - b][AX + b - b]') = E(AXX'A') = A \cdot E(XX') \cdot A$$

and

$$\phi_Y(u) = E\exp(iu'Y) = E\exp(iu'(AX + b)) = E(\exp(iu\prime b) * \exp(iu\prime AX)) = \exp(iu'b) \prod_k E\exp(i(u'A)_k X_k)$$

where $(u'A)_k$ is the k-th component of the vector $u'A$.

and

$$Eexp(iaX_i = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(iax)\exp(-x^2/2)dx = exp(-a^/2)$$

then

$$\phi_Y(u) = \exp(iu'b - \frac{1}{2}u'\Sigma_{YY}u)$$

Definition (Standard Brownian Motion) A standard Brownian motion or a standard Wiener Process $\{B(t)\,,\, t \geq 0\}$ is a stochastic process satisfying:

- (a) $B(0) = 0$
- (b) for every t and $0 = t_0 < t_1 < \cdots < t_n$, $\Delta_k := B(t_k) - B(t_{k-1})$ are independent
- (c) $B(t) - B(s) \sim N(0, t - s)$ for $t \leq s$

(c) is saying $B(t + 1) - B(t) \sim N(0, 1)$ basically

Definition (Poisson Process) A Poisson Process $\{N(t)\,,\, t \leq 0\}$ with a **mean rate** λ is a stochastic process statisfying:

- (a) $N(0) = 0$
- (b) for every t and $0 = t_0 < t_1 < \cdots < t_n$, $\Delta_k := N(t_k) - N(t_{k-1})$ are independent
- (c) $N(t) - N(s) \sim \mathrm{Po}(\lambda(t - s))$ for $t \geq s$

Poisson (a) (b) is then same as SBM (a) (b) the diff is SBM(c) is standard but Pois(c) is a Poisson distribution.

A.2 Hilbert Spaces

Definition (Hilbert Space) A space $\mathcal{H}$ is (complex) Hilbert space if:

- $\mathcal{H}$ is a vector space i.e.
  - (a) $\mathcal{H}$ has an addition $+$ that $\mathcal{H} + \mathcal{H} = \mathcal{H}$
  - (b) $\mathcal{H}$ has a multiplication $\mu : \mathbb{C} \times \mathcal{H} \to \mathcal{H}$ than $\mathbb{C} \cdot \mathcal{H} = \mathcal{H}$
- $\mathcal{H}$ is a inner-product space
  - (a) $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is bi-linear function
  - (b) $||x|| = \sqrt{\langle x, x \rangle}$ is a norm i.e $||x|| \geq 0$ and $||x|| = 0$ iff $x = 0$
- $\mathcal{H}$ is complted i.e. $\{x_k\}_k$ is *cauchy series* then exists $x \in \mathcal{H}$ such that $\lim_{k \to \infty} ||x_k - x|| = 0$

Definition $\mathcal{H}$ is real Hilbert Space if $\mathbb{C}$ is replaced with $\mathbb{R}$

convergence of random variables

- $(X_n \xrightarrow{m.s.} X)$ iff $(||X_n - X|| \to 0)$ and $X\,,\, X_i \in L^2$ (mean-square convergence)
- $(X_n \xrightarrow{P} X)$ iff $P(|X_n - X| > \epsilon) \to 0$ for all $\epsilon > 0$ (convergence in probability)
- $(X_n \xrightarrow{a.s.} X)$ iff $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega - E$ where $P(E) = 0$ (almost sure convergence)/(convergence with probability one)

and we have:

$$X_n \xrightarrow{m.s.} X \Rightarrow X_n \xrightarrow{P} X$$

proof

$X_n \xrightarrow{m.s.} X \Rightarrow \forall k \in \mathbb{N}^+ \exists n_k \in \mathbb{N}^+(n \geq n_k \Rightarrow ||X_n - X|| < \frac{1}{k}) \Rightarrow (\forall \epsilon > 0(\exists n_k \in \mathbb{N}^+ P\epsilon > \frac{1}{k})) \Rightarrow (\forall \epsilon \lim_{n \to \infty P(||X_n - X|| >}$
$\square$

$$X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X$$

proof

$X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega - E$ is to say $P(||X_n - X|| > \epsilon) = P(E) = 0$ when $n \to \infty$ $\square$

proposition if $X_n \xrightarrow{P} Y and X_n \xrightarrow{a.s} Y$ then $X = Y a.s.$

Let $\mathcal{M}$ be a Hilbert sub-space of $\mathcal{H}$ i.e. $\mathcal{M} \subset \mathcal{H}$ and $\mathcal{M}$ is Hilbert space.

Let $\mathcal{M}$ be a subset of $\mathcal{H}$. The orthogonal complement of $\mathcal{M}$ which denoted as $\mathcal{M}^\perp$:

$$\mathcal{M}^\perp = \{y \in \mathcal{H} : \langle y, x \rangle = 0, \forall x \in \mathcal{M}\}$$

Theorem (The Projection Theorem) If $\mathcal{M}$ is a Hilbert sub-space $\mathcal{H}$ and $x \in \mathcal{H}$ then

(i) there is a unique element $\hat{x} \in \mathcal{M}$ such that

$$||x - \hat{x}|| = \inf_{y \in \mathcal{M}} ||x - y||$$

(ii) $\hat{x} \in \mathcal{M} = \inf_{y \in \mathcal{M}||x-y||}$ iff $\hat{x} \in \mathcal{M}$ and $x - \hat{x} \in \mathcal{M}^{\perp}$

$P_{\mathcal{M}}x := \hat{x}$ is called (orthogonal) projection of x onto $\mathcal{M}$

Definition (Closed Spac) The closed span $\bar{X}$ of X a subset of hilbert space is defined to be the smallest Hilber sub-space which contains X ,it is to say that for any $\mathcal{M}$ is a Hilbert sub-space of $\mathcal{H}$ that exists a Hilbert sub-space $\bar{X}$ satisfying
$X \subset \mathcal{M} \Rightarrow \bar{X} \subset \mathcal{M}$

Properties (Properties of Projections) Let $\mathcal{H}$ be a Hilbert space and $P_{\mathcal{M}}$ be the projection then:

(i) $P_{\mathcal{M}}(\alpha x + \beta y) = \alpha P_{\mathcal{M}}x + \beta P_{\mathcal{M}}y \quad x\,,\, y \in \mathcal{H}, \quad \alpha\,,\, \beta \in \mathbb{C}/\mathbb{R}$
(ii) $||x||^2 = ||P_{\mathcal{M}}x||^2 + ||(I - P_{\mathcal{M}}x)||^2$ where I is the identity mapping
(iii) $x = P_{\mathcal{M}}x + (I - P_{\mathcal{M}}x)$