STA457 Notes

Ian Zhang

1 Characteristics of Time Series

Definition (Time series)**.** A time series is a set of observations x_t, each recorded at time t.

If the set of times T_0 is discrete, then the time series is discrete.

Definition (White noise)**.** White noise is a time series $\{w_t\}$ where the w_t are uncorrelated and $E(w_t) = 0$, $\text{Var}(w_t) = \sigma_w^2 < \infty$. We denote white noise by $w_t \sim wn(0, \sigma_w^2)$.

Some examples of time series are

- Moving average: used to "smooth" a series

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$$

  is a moving average of order 3

- Autoregressive model: the output variable of an autoregressive model depends on past values of the series, for example

$$x_t = x_{t-1} = 0.9x_{t-2} + w_t$$

  where w_t is white noise

- Random walk with drift:

$$x_t = \delta + x_{t-1} + w_t$$

  with initial condition $x_0 = 0$, δ is the drift, $w_t \sim wn(0, \sigma_w^2)$.

Definition (Autocovariance function)**.** The autocovariance function of a time series x_t is defined as

$$\gamma_x(s,t) = \text{Cov}(x_s, x_t)$$

If $U = \sum_{i=1}^n a_i X_i$ and $V = \sum_{j=1}^m b_j Y_j$, then

$$\text{Cov}(U,V) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(X_i, Y_j)$$

Definition (Autocorrelation function (ACF))**.** The ACF of a time series x_t is

$$\rho_x(s,t) = \frac{\gamma_x(s,t)}{\sqrt{\gamma_x(s,s)\gamma_x(t,t)}}$$

We can also look at the covariance and correlation functions of two time series x_t and y_t.

Definition (Cross-covariance function)**.** The cross-covariance function of x_t and y_t is

$$\gamma_{xy}(s,t) = \text{Cov}(x_s, y_t)$$

Definition (Cross-correlation function (CCF))**.** The CCF of x_t and y_t is

$$\rho_{xy}(s,t) = \frac{\gamma_{xy}(s,t)}{\sqrt{\gamma_x(s,s)\gamma_y(t,t)}}$$

1.1 Stationary Time Series

Definition (Weakly staionary time series)**.** A time series w_t is weakly stationary if the following hold:

(i) the mean function, μ_t, is constant and independent of t

(ii) the autocovariance function, $\gamma_x(s,t)$, depends on s and t only through their absolute difference $|s-t|$

We let *stationary* mean *weakly stationary.*

For autocovariance function, we can let $s = t + h$, so $\gamma_x(s,t) = \gamma_x(h)$. If x_t is stationary, then $\gamma_x(h) = \phi(|h|)$ where ϕ is some function of $|h|$. In this form, the ACF of a stationary time series is

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

1.1.1 Jointly Stationary Time Series

Definition (Jointly stationary)**.** Two time series x_t and y_t are jointly stationary if x_t and y_t are each respectively stationary and the cross-covariance function $\gamma_{xy}(h) = \text{Cov}(x_{t+h}, y_t)$ is a function only of lag h.

Definition (CCF of jointly stationary)**.** The CCF of jointly stationary time series x_t and y_t is

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}}$$

1.2 Estimation of Correlation

If a time series is stationary, then its mean is a constant μ, thus we can estimate it by the sample mean

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

The variance of sample mean is

$$\begin{aligned}
\text{Var}(\bar{x}) &= \text{Var}\left(\frac{1}{n}\sum_{t=1}^{n} x_t\right) \\
&= \frac{1}{n^2}\text{Cov}\left(\sum_{t=1}^{n} x_t, \sum_{s=1}^{n} x_s\right) \\
&= \frac{1}{n}\sum_{h=-n}^{n}\left(1 - \frac{|h|}{n}\right)\gamma_x(h)
\end{aligned}$$

The sample autocovariance is

$$\hat{\gamma}_x(h) = \frac{1}{n}\sum_{t=1}^{n=h}(x_{t+h} - \bar{x})(x_t - \bar{x})$$

with $\hat{\gamma}_x(h) = \hat{\gamma}_x(-h)$ for $h = 0, 1, \ldots, n-1$. The sample ACF is thus

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

2 Time Series Regression

Consider the MLR model

$$x_t = \beta_0 + \beta_1 z_{t1} + \cdots + \beta_q z_{tq} + w_t$$

where $w_t \sim \mathcal{N}(0, \sigma_w^2)$ for $t = 1, \ldots, n$. Let $z_t = (1, z_{t1}, \ldots, z_{tq})^T$, $\beta = (\beta_0, \beta_1, \ldots, \beta_q)^T$, thus we can rewrite the model as

$$x_t = \beta^T z_t + w_t$$

2.1 Ordinary Least Squares Estimation

OLS estimation finds $\hat{\beta}$ that minimizes

$$Q = \sum_{t=1}^{n} w_t^2 = \sum_{t=1}^{n}(x_t - \beta^T z_t)^2$$

If the matrix $\sum_{t=1}^{n} z_t z_t^T$ is non-singular, then the LSE of β is

$$\hat{\beta} = \left(\sum_{t=1}^{n} z_t z_t^T\right)^{-1}\sum_{t=1}^{n} z_t x_t$$

Some properties of $\hat{\beta}$ are

- $E(\hat{\beta}) = \beta$
- $\hat{\beta} \sim \mathcal{N}_p(\beta, \sigma_w^2 C)$ where $p = q + 1$, $C = \left(\sum_{t=1}^{n} z_t z_t^T\right)^{-1}$

If $\hat{\sigma}_w^2$ is unknown, we can estimate using

$$s_w^2 = MSE = \frac{SSE}{n-p}$$

where

$$SSE = \sum_{t=0}^{n}(x_t - \hat{\beta}^T z_t)^2$$

There are also a few tests we can consider. If we want to test each β_i individually in $H_0 : \beta_i = 0$, then consider the test statistic

$$t = \frac{\hat{\beta}_i - \beta_i}{s_w\sqrt{c_{ii}}} \sim t_{n-p}$$

where c_{ii} is the ith element on the diagonal of C.
If we want to test whether only a subset of $r < q$ independent variables, $z_{t,1:r} = \{z_{t1}, \ldots, z_{tr}\}$ is influencing x_t (i.e.: $\beta_{r+1}, \ldots, \beta_q = 0$), we use the test statistic

$$F = \frac{(SSE_r - SSE)/(q-r)}{SSE/(n-p)} = \frac{SSR/(q-r)}{SSE/(n-p)} = \frac{MSR}{MSE} \sim F_{n-p}^{q-r}$$

We reject at level α if $F_c > F_{n-p}^{q-r}(\alpha)$ where F_c is the value of test statistic under H_0.
Suppose we have a model with k coefficients (smaller model). Then, the maximum likelihood estimator for σ_k^2 is

$$\hat{\sigma}_k^2 = \frac{SSE(k)}{n}$$

where $SSE(K)$ is the SSE under the smaller model. Define Akaike's Information Criteria (AIC) by

$$AIC = \log(\hat{\sigma}_k^2) + \frac{n+2k}{n}$$

The value of k yielding the smallest AIC indicates the best model.

3 Exploratory Data Analysis

Definition. The trend stationary model can be written as

$$x_t = \mu_t + y_t$$

where μ_t is the observed trend and y_t is a stationary process.

Denote the first difference by

$$\nabla x_t = x_t - x_{t-1}$$

Definition (Backshift operator)**.** The backshift operator is

$$Bx_t = x_{t-1}$$

We can extend this definition to powers: since $Bx_{t-1} = x_{t-2}$ but $x_{t-1} = Bx_t$, then we have

$B^2 x_t = x_{t-2}$. By induction, we have $B^k x_t = x_{t-k}$.

Definition (Differences)**.** Differences of order d are

$$\nabla^d = (1 - B)^d$$

4 ARIMA Models

4.1 AR(p) Models

Definition (AR(p) model)**.** An autoregressive model of order p, denoted AR(p), is of the form

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t$$

where x_t is stationary, $w_t \sim wn(0, \sigma_w^2)$, $\phi_1, \ldots, \phi_p$ are constants with $\phi_p \neq 0$.

If the mean μ of x_t is non-zero, then we can replace x_t with $x_t - \mu$, so

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + w_t$$
$$\iff x_t = \alpha + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t$$

where $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$. Using the backshift operator,

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) x_t = w_t \implies \phi(B) x_t = w_t$$

where $\phi_B = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$.

Proposition. Let $x_t = \phi x_{t-1} + w_t$ be an AR(1) process where $|\phi| < 1$ and $\sup_t \operatorname{Var}(x_t) < \infty$. Then

- $x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}$. That is, x_t is a linear process
- The autovariance function $\gamma(h)$ is
$$\gamma(h) = \frac{\sigma_w^2 \phi^h}{1 - \phi^2}$$
- The autocorrelation function is
$$\rho(h) = \phi^h$$

Proof. To show x_t is a linear process, by recursion we have

$$\begin{aligned} x_t &= \phi x_{t-1} + w_t \\ &= \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \cdots \\ &= \phi^k x_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j} \end{aligned}$$

and so on.
For the autocovariance,

$$\begin{aligned}\gamma(h) &= \text{Cov}(x_{t+h}, x_t)\\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\phi^i\phi^j\text{Cov}(w_{t+h-i}, w_{t-j})\\ &= \sigma_w^2\sum_{j=0}^{\infty}\phi^{j+h}\phi^j\\ &= \sigma_w^2\phi^2\sum_{j=0}^{\infty}\phi^{2j}\\ &= \frac{\sigma_w^2\phi^h}{1-\phi^2} && \text{since } |\phi|< 1\end{aligned}$$

For the ACF, since x_t is stationary and $\sup_t \text{Var}(x_t) < \infty$, then

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\frac{\sigma_w^2\phi^h}{1-\phi^2}}{\frac{\sigma_w^2}{1-\phi^2}} = \phi^h$$

as required. ■

4.1.1 Explosive Models and Causality

Consider the simple walk $x_t = x_{t-1} + w_t$, $w_t \sim w_n(0, \sigma_w^2)$. Since

$$\gamma(s,t) = \min\{s,t\}\sigma_w^2$$

x_t is not stationary.
An AR(1) process $x_t = \phi x_{t-1} + w_t$ with $|\phi|> 1$ is *explosive* since its values will increase quickly. Since $|\phi|^{-1}< 1$, this suggests the stationary future dependent AR(1) model

$$x_t = -\sum_{j=1}^{\infty}\phi^{-j}w_{t+j}$$

4.2 MA(q) Models

Definition (MA(q) models)**.** The moving average model of order q, MA(q), is

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

where $w_t \sim wn(0, \sigma_w^2)$, $\theta_1, \ldots, \theta_q$ are constants where $\theta_q \neq 0$.

We can rewrite $x_t = \theta(B)w_t$ where $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ is the moving average operator.

4.2.1 Invertibility

If $|\theta|< 1$, then MA(1) as $w_t = -\theta w_{t-1}$ can be written as

$$w_t = \sum_{j=0}^{\infty} (-\theta)^j x_{t-j}$$

4.3 ARMA(p, q) models

Definition. A time series $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ is ARMA(p, q) if it's stationary and

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

with $\phi_p, \theta_q \neq 0$, $\sigma_w^2 > 0$.

If x_t has nonzero mean μ, then set $\alpha = \mu(1 - \phi_1 - \phi_2 - \cdots - \phi_p)$ and write the model as

$$x_t = \alpha + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

where $w_t \sim wn(0, \sigma_w^2)$.

4.3.1 Parameter redundancy

Consider $x_t = 0.5x_{t-1} - 0.5w_{t-1} + w_t$, which looks ARMA$(1, 1)$. However,this is equivalent to

$$(1 - 0.5B)x_t = (1 - 0.5B)w_t \iff x_t = w_t$$

so x_t is just white noise. This is due to paramter redundancy.

Definition. The AR and MA polynomials are respectively defined as

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p, \phi_p \neq 0$$

$$\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q, \theta_q \neq 0$$

where $z \in \mathbb{C}$.

4.4 Causality and Invertibility

Definition (Causal model)**.** An ARMA(p, q) model is causal if $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ can be written as a one-sided linear process:

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B) w_t$$

where $\psi(B) := \sum_{j=0}^{\infty} \psi_j B^j$, $\sum_{j=0}^{\infty} |\psi_j| < \infty$. We set $\psi_0 = 1$.

Definition (Invertible model)**.** An ARMA(p,q) model is invertible if $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ can be written as

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t$$

where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$, $\sum_{j=0}^{\infty} |\pi_j| < \infty$. Set $\pi_0 = 1$.

Proposition. An ARMA(p,q) model is causal iff $\phi(z) \neq 0$ for $|z| \leq 1$. The coefficients ψ_j of the linear process can be determined by solving

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, |z| \leq 1$$

An ARMA(p,q) process is causal only when the roots of $\phi(z)$ lie outside the unit circle (i.e.: $\phi(z) = 0$ only if $|z| > 1$).

Proposition. An ARMA(p,q) model is invertible iff $\theta(z) \neq 0$ for $|z| \leq 1$. The coefficients π_j can be solved through

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, |z| \leq 1$$

An ARMA(p,q) process is invertible only when roots of $\theta(z)$ lie outside unit circle (i.e.: $\theta(z) = 0$ only if $|z| > 1$).

4.4.1 Partial ACF

For a MA(q) model, the ACF is 0 for lag values greater than h. Since $\theta_q \neq 0$, then the ACF will not be 0 at lag q, thus we can use this to identify MA(q) models. For AR models, however, use the partial ACF (PACF) to identify.

Definition. Suppose X, Y and Z are random variables. By regression X on Z to obtain $\hat{X}$ and Y on Z to obtain $\hat{Y}$, the PACF of X and Y given Z is

$$\rho_{X,Y|Z} = \text{Corr}(X - \hat{X}, Y - \hat{Y})$$

Let x_t be stationary with mean 0. For $h \geq 2$, let $\hat{x}_{t+h}$ denote the regression of x_{t+h} on $\{x_{t+h-1}, x_{t+h-2}, \dots, x_{t+1}\}$:

$$\hat{x}_{t+h} = \beta_1 x_{t+h-1} + \beta_2 x_{t+h-2} + \cdots + \beta_{h-1} x_{t+1}$$

Let $\hat{x}_t$ denote the regression of x_t on $\{x_{t+1}, \dots, x_{t+h-1}\}$:

$$\hat{x}_t = \beta_1 x_{t+1} + \cdots + \beta_{h-1} x_{t+h-1}$$

Since x_t is stationary, then all the β_k are the same in both regressions (i.e.: $\beta_k = \beta_k'$). So, the PACF of x_t, denoted ϕ_{hh}, is

$$\phi_{hh} = \text{Corr}(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t)$$

with $\phi_{11} = \text{Corr}(x_t + 1, x_t) = \rho(1)$.
The PACF represents the correlation between x_{t+h} and x_t with linear dependence removed.

4.5 Forecasting

Goal: To predict future values of a time series x_{n+m} for $m = 1, 2, \ldots$ based on colleted data $x_{1:n} = \{x_1, \ldots, x_n\}$.

- For our purposes, assume a stationary x_t

Definition. The minimum MSE predictor of x_{n+m} is

$$x_{n+m}^n = E(x_{n+m} \mid x_{1:n})$$

since conditional expectation minimizes $E[x_{n+m} - g(x_{1:n})]^2$ over all functions g of $x_{1:n}$.

- For our purposes, we only consider linear predictors:

$$x_{n+m}^n = \alpha_0 + \sum_{k=1}^{n} \alpha_k x_k$$

  This is called the best linear predictor (BLP).

Given data $x_1, \ldots, x_n$, the BLP coefficients in

$$x_{n+m}^n = \alpha_0 + \sum_{k=1}^{\infty} \alpha_k x_k, m \geq 1$$

are found by solving

$$E[(x_{n+m} - x_{n+m}^n)x_k] = 0$$

for all $k = 0, 1, \ldots, n$. We set $x_0 = 1$.

4.5.1 Prediction Error

Definition. The mean squared m-step-ahead prediction error is

$$P_{m+n}^n = E[x_{n+m} - x_{n+m}^n]^2$$

We use prediction intervals to assess precision of forecasts. These intervals are given by

$$x_{n+m}^n \pm C_{\alpha/2}\sqrt{P_{n+m}^n}$$

4.6 Estimation

Assume observations $x_1, \ldots, x_n$ are from a causal and invertible Gaussian ARMA(p, q) model. The goal is to estimate the parameters of the model and σ_w^2.

Consider an AR(p) $x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t$. The Yule-Walker equations are

$$\begin{aligned}\gamma(h) &= \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \cdots + \phi_p\gamma(h-p) \\ \sigma_w^2 &= \gamma(0) - \phi_1\gamma(1) - \cdots - \phi_p\gamma(p)\end{aligned}$$

In matrix notation,

$$\Gamma_p\phi = \gamma_p \qquad \sigma_w^2 = \phi(0) - \phi'\gamma_p$$

where $\Gamma_p = (\gamma(k-j))_{j,k=1}^p$ is a $p \times p$ matrix and $\phi = (\phi_1, \ldots, \phi_p)^T$ and $\gamma_p = (\gamma(1), \ldots, \gamma(p))^T$. We estimate using Method of Moments:

$$\hat{\phi} = \hat{\Gamma}_p^{-1}\hat{\gamma}(h) \qquad \hat{\sigma}_w^2 = \hat{\gamma}(0) - \hat{\gamma}_p^T\hat{\Gamma}_p^{-1}\hat{\gamma}_p$$

We can also use the sample ACF:

$$\hat{\phi} = \hat{R}_p^{-1}\hat{\rho}_p \qquad \hat{\sigma}_w^2 = \hat{\gamma}(0)[1 - \hat{\rho}_p^T\hat{R}_p^{-1}\hat{\rho}_p]$$

The asymptotic behaviour of Yule-Walker estimators for a causal AR(p) is

$$\begin{gathered}\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{d} \mathcal{N}(0, \sigma_w^2\Gamma_p^{-1}) \\ \hat{\sigma}_w^2 \xrightarrow{p} \sigma_w^2\end{gathered}$$

So, the PACF satisfies

$$\sqrt{n}\hat{\phi}_{hh} \xrightarrow{d} \mathcal{N}(0,1)$$

for $h > p$.

4.7 ARIMA Models

Definition. A process x_t is ARIMA(p, d, q) if

$$\nabla^d x_t = (1-B)^d x_t$$

is ARMA(p, q). In general,

$$\phi(B)(1-B)^d x_t = \theta(B) w_t$$

If $E(\nabla^d x_t) = \mu$, then

$$\phi(B)(1-B)^d x_t = \delta + \theta(B) w_t$$

where $\delta = \mu(1 - \phi_1 - \cdots - \phi_p)$.

4.7.1 Fitting ARIMA Models

Steps:

1. Plot the data

2. Transform the data if needed
3. Identifying dependence orders of model from ACF/PACF
4. Parameter estimation
5. Diagnostics
   - Standardized residuals computed by
     $$e_t = \frac{x_t - \hat{x}_t^{t-1}}{\sqrt{\hat{P}_t^{t-1}}}$$
     where $\hat{x}_t^{t-1}$ is the one-step-ahead prediction of x_t based on the fitted model and $\hat{P}_t^{t-1}$ is the prediction error
6. Model choice

For regression with autocorrelated errors, the model looks like

$$y_t = \sum_{j=1}^{r} \beta_j z_{tj} + x_t$$

where x_t is a process with covariance function $\gamma_x(s,t)$. To identify the model, these are the steps:

1. Run regular regression of y_t on $z_{t1}, \dots, z_{tr}$ and retain the residuals
2. Identify an ARMA model for the residuals $\hat{x}_t$ following the steps above
3. Run weighted least squares on regression model with autocorrelated errors using model in step 2
4. Inspect residuals $\hat{w}_t$

4.7.2 Multiplicative Seasonal ARMIA Models

Definition. The pure seasonal ARMA, denoted ARMA$(P,Q)_s$ is

$$\Phi_P(B^s)x_t = \Theta_Q(B^s)w_t$$

where

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}$$
$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Qs}$$

are respectively the seasonal autogressive and moving average operators of orders P and Q with period s.

Similary to an ARMA(p,q) model, this model is causal/invertible only if the roots of $\Phi_P(z)$/ $\Theta_Q(z)$ lie outside the unit circle. We estimate the orders P and Q with the chart

	$\text{AR}(P)_s$	$\text{MA}(Q)_s$	$\text{ARMA}(P,Q)_s$
ACF*	Tails off at lags ks, $k = 1, 2, \cdots,$	Cuts off after lag Qs	Tails off at lags ks
PACF* *	Cuts off after lag Ps	Tails off at lags ks $k = 1, 2, \cdots,$	Tails off lags ks

This chart can also be used to identify the order of regular ARMA models, albeit without the s. In general, seasonal and nonseasonal operators are combined into a model $\text{ARMA}(p,q) \times (P,Q)_s$, given by

$$\Phi_P(B^s)\phi(B)x_t = \Theta_Q(B^s)\theta(B)w_t$$

Definition. A SARIMA model $\text{ARIMA}(p,d,q) \times (P,D,Q)_s$ is given by

$$\Phi_P(B^s)\phi(B)\nabla_s^D\nabla_d x_t = \delta + \Theta_Q(B^s)\theta(B)w_t$$

where $\nabla_s^D = (1 - B^s)^D$.

5 Spectal Analyis and Filtering

Consider a series

$$x_t = \sum_{k=1}^{q}[U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t)]$$

where U_{k1} and U_{k2} are uncorrelated with mean 0, variance σ_k^2. The ACF of this series is

$$\gamma_x(h) = \sum_{k=1}^{q}\sigma_k^2\cos(2\pi\omega_k h)$$

5.1 Estimation

For any time series sample $x_1, \ldots, x_n$, if n is odd, then we can write

$$x_t = a_0 + \sum_{j=1}^{\frac{n-1}{2}}[a_j\cos(2\pi tj/n) + b_j\sin(2\pi tj/n)]$$

for $t = 1, \ldots, n$. If n is even, then

$$x_t = a_0 + \sum_{j=1}^{\frac{n}{2}}[a_j\cos(2\pi tj/n) + b_j\sin(2\pi tj/n)]$$

For the coefficients, choose $a_0 = \bar{x}$ and

$$a_j = \frac{2}{n}\sum_{t=1}^{n}x_t\cos(2\pi tj/n)$$

$$b_j = \frac{2}{n}\sum_{t=1}^{n} x_t \sin(2\pi tj/n)$$

Definition. The scaled periodigram is $P(j \mid n) = a_j^2 + b_j^2$.

This function indicates which frequency components are large in magnitude and which are small. It serves as an estimate of σ_j^2 corresponding to $\omega_j = \frac{j}{n}$.

Definition. The discrete Fourier transform is

$$d(j \mid n) = n^{-\frac{1}{2}} \left(\sum_{t=1}^{n} x_t \cos(2\pi tj/n) - i \sum_{t=1}^{n} x_t \sin(2\pi tj/n) \right)$$

Rewriting in Euler form,

$$d(j \mid n) = n^{-\frac{1}{2}} \sum_{t=1}^{n} x_t e^{-i2\pi tj/n}$$

The periodogram is

$$|d(j \mid n)|^2 = \frac{1}{n}\left(\sum_{t=1}^{n} x_t \cos(2\pi tj/n) \right)^2 + \frac{1}{n}\left(\sum_{t=1}^{n} x_t \sin(2\pi tj/n) \right)^2$$

Notice that the scaled periodogram is $P(j \mid n) = \frac{4}{n}|d(j \mid n)|^2$.

5.2 The Spectral Density

If $\{x_t\}_t$ is a stationary time series with autocovariance $\gamma(h)$, then there exists a unique monotically increasing function $F(\omega)$ with $F(-\infty) = F(-\frac{1}{2}) = 0$, $F(\infty) = F(\frac{1}{2}) = \gamma(0)$ such that

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i\omega h} dF(\omega)$$

If $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, then it has the reprsentation

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega) e^{2\pi i\omega h} d\omega$$

as the *inverse Fourier transform* of the spectral density

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h}, -\frac{1}{2} \leq \omega \leq \frac{1}{2}$$

The spectral density **behaves like any density function:**

- $f(\omega) \geq 0$
- $f(\omega) = f(-\omega)$

For $h = 0$, we have that

$$\gamma(0) = \text{Var}(x_t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega) d\omega$$

Some properties of the spectral density:

- If $f(\omega), g(\omega)$ are spectral densities such that
$$\gamma_f(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega)e^{-2\pi i\omega h}d\omega = \int_{-\frac{1}{2}}^{\frac{1}{2}} g(\omega)e^{2\pi i\omega h}d\omega = \gamma_g(h)$$
then the two spectral densities are equal (i.e. the spectral density is **unique**)

- If $y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j}$ where $\sum_{j=-\infty}^{\infty}|a_j|< \infty$, let $A(\omega) = \sum_{j=-\infty}^{\infty} a_j e^{-2\pi i\omega j}$. Then $f_y(\omega) = |A(\omega)|^2 f_x(\omega)$

- If x_t is ARMA(p, q) with autoregressive operator $\phi(z)$ and moving average operator $\theta(z)$, then
$$f_x(\omega) = \sigma_w^2 \frac{|\theta(e^{-2\pi i\omega})|^2}{|\phi(e^{-2\pi i\omega})|^2}$$

We can relate this to the periodogram. Let $\omega_j = \frac{j}{n}$. Then the DFT is
$$d(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^{n} x_t e^{-2\pi i\omega_j t}$$

Solve for x_t using the inverse DFT
$$x_t = n^{-\frac{1}{2}} \sum_{j=0}^{n-1} d(\omega_j)e^{2\pi i\omega_j t}$$

The periodogram is $I(\omega_j) = |d(\omega_j)|^2$ and is the sample version of $f(\omega_j)$. Thus the periodogram can be viewed as the sample spectral density.
If $x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$ with $\sum_{j=-\infty}^{\infty}|\psi_j|< \infty$ where $w_t \sim wn(0, \sigma_w^2)$, then for any collection of m distinct frequencies $\omega_j \in (0, \frac{1}{2})$ with $\omega_{j:n} \to \omega_j$, then
$$\frac{2I(\omega_{j:n})}{f(\omega_j)} \xrightarrow{d} \chi_2^2$$

provided $f(\omega_j)$ for $j = 1, \ldots, m$. Thus, a $100(1-\alpha)\%$ confidence interval for $f(\omega)$ is
$$\frac{2I(\omega_{j:n})}{\chi_2^2(1-\frac{\alpha}{2})} \leq f(\omega) \leq \frac{2I(\omega_{j:n})}{\chi_2^2(\frac{\alpha}{2})}$$

5.3 Non-parametric estimation

For $\omega^* = \omega_j + \frac{k}{n}$, let
$$\mathbb{B} = \left\{\omega^* : \omega_j - \frac{m}{n} \leq \omega^* \leq \omega_j + \frac{m}{n}\right\}$$
where $L = 2m + 1$ is odd, chosen such that
$$f\left(\omega_j + \frac{k}{n}\right) \approx f(\omega)$$

for $-m \leq k \leq m$.

Definition. The smoothed periodogram is defined as

$$\bar{f}(\omega) = \frac{1}{L} \sum_{k=-m}^{m} I\left(\omega_j + \frac{k}{n}\right)$$

This function satisfies

$$\frac{2L\bar{f}(\omega)}{f(\omega)} \sim \chi^2_{2L}$$

so a $100(1-\alpha)\%$confidence interval of $f(\omega)$ is

$$\frac{2L\bar{f}(\omega)}{\chi^2_{2L}(1-\frac{\alpha}{2})} \leq f(\omega) \leq \frac{2L\bar{f}(\omega)}{\chi^2_{2L}(\frac{\alpha}{2})}$$

Definition. A parametric spectral estimator is obtained by fitting an AR(p) where the order p is determined by AIC. If $\hat{\phi}_1, \ldots, \hat{\phi}_p$ and σ^2_w are the fitted AR(p) to x_t, then

$$\hat{f}_x(\omega) = \frac{\sigma^2_w}{|\hat{\phi}(e^{-2\pi i\omega})|^2}$$

where $\hat{\phi}(z) = 1 - \sum_{j=1}^{p} \hat{\phi}_j z^j$.

A $100(1-\alpha)\%$ confidence interval of $f_x(\omega)$ is

$$\frac{\hat{f}_x(\omega)}{1 + Cz_{\alpha/2}} \leq f_x(\omega) \leq \frac{\hat{f}_x(\omega)}{1 - Cz_{\alpha/2}}$$

where $C = \sqrt{\frac{2p}{n}}$ and $z_{\alpha/2}$ is the $\alpha/2$ quantile of $\mathcal{N}(0,1)$.

5.4 Multiple series and cross-spectra

The cross-covariance of jointly stationary x_t and y_t has the representation

$$\gamma_{xy}(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_{xy}(\omega) e^{2\pi i\omega h} d\omega$$

for $h = 0, \pm 1, \pm 2, \ldots$ where $f_{xy}(\omega)$ is the cross-spectrum defined as

$$f_{xy}(\omega) = \sum_{h=-\infty}^{\infty} \gamma_{xy}(h) e^{-2\pi i\omega h}$$

We can rewrite $f_{xy}(\omega)$ in polar form:

$$f_{xy}(\omega) = \sum_{h=-\infty}^{\infty} \gamma_{xy}(h)\cos(2\pi\omega h) - i \sum_{h=-\infty}^{\infty} \gamma_{xy}(h)\sin(2\pi\omega h)$$

Since $\gamma_{xy}(h) = \gamma_{yx}(-h)$, then $f_{xy}(\omega) = f^*_{yx}(\omega)$ where $*$ denotes the conjugate of a complex number.

Definition. The squared coherence function is

$$\rho_{yx}^2 = \frac{|f_{yx}(\omega)|^2}{f_{xx}^2(\omega) f_{yy}^2(\omega)}$$