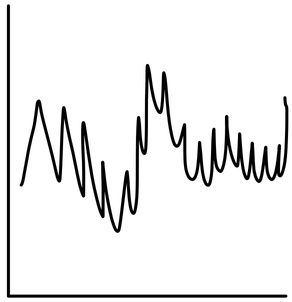
MAT 3379

Time Series Analysis

Study Guide

Winter 2024



$$X_t = Y_t + m_t + S_t$$

Examples

- White noise: $\{Z_t\}$ sequence of independent random variables with mean 0 and variance 1
- Random Walk: $\{Z_t\}$ sequence of iid random variables with mean 0 and variance σ_Z^2. $X_t = \sum_{i=1}^{t} Z_t$, $t=1,2,\ldots$
- Model with trend: Sometimes a trend is present in time series. $X_t = 1+2t+Z_t$, $t=1,2,\ldots$ Trend is $m_t = 1+2t$
- Economics Trend: $X_t = P_t e^{rt}$ P_t is real price, r is interest rate

Time Series Structure

$$X_t = m_t + Y_t + S_t$$

- m_t is a trend
- S_t is a seasonal part
- Y_t is a stationary part

Eliminate Trend

- Differencing: Compute $\nabla X_t = X_t - X_{t-1}$, $t=2,\ldots,n$
- Polynomial Fitting: Assume $m_t = a+bt$. Estimate a and b by minimizing $\sum_{t=1}^{n}(X_t - a - bt)^2$, $\hat{m}_t = \hat{a} + \hat{b}t$. Detrended time series: $\hat{Y} = X_t - \hat{m}_t$
- Exponential Smoothing: $\alpha \in (0,1)$ Trend: $\hat{m}_1 = X_1$, $\hat{m}_t = \alpha X_t + (1-\alpha)\hat{m}_{t-1}$, $t=2,\ldots,n$ De-trended time series: $\hat{Y}_t = X_t - \hat{m}_t$
- Moving Average Smoothing: $q \in \mathbb{Z}^+$ $\hat{m}_t = (2q+1)^{-1}\sum_{j=-q}^{q} X_{t+j}$ $q+1 \le t \le n-q$ Detrended time series: $\hat{Y}_t = X_t - \hat{m}_t$

Mean Function

- $\mu_X(t) = E[X_t]$

Covariance Function

- $\gamma_X(t,s) = Cov(X_t, X_s) = E[X_tX_s] - E[X_t]E[X_s]$

Note: $\gamma_X(t,t) = Var(X_t)$

Properties of Covariance

- For $a \in \mathbb{R}$, $Cov(X,a) = 0$
- For $a,b \in \mathbb{R}$, $Cov(X, aU+bV) = a\,Cov(X,U) + b\,Cov(X,V)$
- $Cov(X,Y)^2 \le Var(X)Var(Y)$

Stationary Time Series

- $\mu_X(t)$ does not depend on t
- $\gamma_X(t,s)$ depends only on $h=t-s$
- Covariance function is non-negative definite

Some useful Properties

- $Cov(A,B) = E[AB] - E[A]E[B]$
- If A, B are independent, then $Cov(A,B) = 0$
- $E[aA+bB] = aE[A] + bE[B]$
- $Cov(A,A) = Var(A)$
- $Var(A+bB) = Var(A) + b^2 Var(B)$ if A, B are independent

MA(1) Model

$\{Z_t\}$ white noise

$\theta \in \mathbb{R}$ $(\theta \neq 0)$

$X_t = Z_t + \theta Z_{t-1}$

- $\mu_X(t) = 0$
- $\gamma_X(t,t+h) = \begin{cases} (1+\theta^2)\sigma_Z^2; & h=0 \\ \theta\sigma_Z^2; & h=1 \\ 0; & h \ge 2 \end{cases}$

Autocorrelation

- $\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{Cov(X_0,X_h)}{Var(X_0)}$

Partial Autocorrelation

- Correlation between X_t and X_{t+h} after conditioning out the "in-between" variables $X_{t+1},\ldots,X_{t+h-1}$

PACF between X_1 and X_2 when conditioning out X_3:

$$\rho_{12.3} = \frac{Corr(X_1,X_2) - Corr(X_1,X_3)Corr(X_2,X_3)}{\sqrt{1-Corr^2(X_1,X_3)}\sqrt{1-Corr^2(X_2,X_3)}}$$

Sample Mean, Sample Autocovariance, Sample Autocorrelation

- Sample Mean: $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Sample Variance: $\hat{\sigma}_X^2 = \hat{\gamma}_X(0) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
- Sample Autocovariance: $\hat{\gamma}_X(h) = \frac{1}{n-1}\sum_{t=1}^{n-h}(X_t - \bar{X})(X_{t+h} - \bar{X})$
- Sample Autocorrelation: $\hat{\rho}_X(h) = \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}$
- Sample PACF at lag 2: $\hat{\alpha}(2) = \frac{\hat{\rho}_X(2) - \hat{\rho}_X(1)}{1 - \hat{\rho}_X^2(1)}$

Linear Processes

Let $\{Z_t\}$ be a white noise

Let $\{\psi_j\}$ be a sequence of constants

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad t = 1,2,\ldots$$

The model is called:

- linear model
- (infinite order) moving average
- (causal) moving average
- (one-sided) moving average

The linear process is well-defined if $\sum_{j=0}^{\infty} |\psi_j| < \infty$

The linear process is also stationary with

- $E[X_t] = 0$
- $\gamma_X(h) = \sigma_Z^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$

Difference and Backward operators

- Difference operator ∇ defined as $\nabla X_t = X_t - X_{t-1}$
- Backward operator B defined as $BX_t = X_{t-1}$

ARMA models

Let $\{Z_t\}$ be a white noise. Then ARMA(p,q) mode is a stationary solution to

$$\phi(B)X_t = \theta(B)Z_t \quad (*)$$

where

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p$$

$$\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q$$

- $\phi(z)$ autoregressive polynomial
- $\theta(z)$ moving average polynomial
- $(*)$ does not guarantee a stationary solution or a causal solution

Stationarity and Causality

The existence of causal and stationary solution means that ARMA model can be written as

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

How to check if there is a stationary and causal solution?

- Take AR polynomial $\phi(z)$
- Find roots of ϕ: $\phi(z) = 0$
- If all roots are such that $|z| \neq 1$, then the model is stationary
- If all roots are such that $|z| > 1$, then the model is causal

Linear Representations

Given an ARMA model. If it is stationary and causal, then we can write it as

$$X_t = \sum_{j=1}^{\infty} \psi_j Z_{t-j}$$

The linear representation of the AR(1) model is

$$X_t = \sum_{j=1}^{\infty} \phi^j Z_{t-j} \quad |\phi| < 1$$

The linear representation of the ARMA(1,1) model is

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

where $\psi_0 = 1$ and $\psi_j = \phi^{j-1}(\theta + \phi)$

Recursive approach to ACF

To find a recursive definition for the covariance of an AR(p) model, first multiply every term by X_{t-h} in the model. Then apply expected value to all terms and use that to find the recursive definition and initial cases for the covariance.

For AR(1), the ACF is

$$\gamma_X(0) = \sigma_Z^2 \frac{1}{1-\phi^2}, \quad \gamma_X(h) = \phi(h-1) \quad h \ge 1$$

ACF and PACF of ARMA(p,q) models

For AR(p) models, the autocorrelation function is never 0 for all lags h. The partial autocorrelation function is 0 for all $h > p$.

For MA(q) models, the autocorrelation function is 0 for all $h > q$. The partial autocorrelation function is never 0.

Yule-Walker Formula for Forecasting

Consider a stationary sequence with $\mu_X = E[X_t]$ and autocovariance $\gamma_X(h)$. We denote $P_n X_{n+k}$, the predicted value of X_{n+k} given that we have n observations $X_1, \ldots, X_n$.

We use linear predictors:

$$P_n X_{n+k} = a_0 + \sum_{i=1}^{n} a_i X_{n+1-i}$$

We minimize the mean square error

$$E\left[(X_{n+k} - P_n X_{n+k})^2\right]$$

Goal: estimate $a_0, a_1, \ldots, a_n$

We obtain $a_0 = \mu\left(1 - \sum_{i=1}^{n} a_i\right)$

If we assume $\mu = 0$, then $a_0 = 0$.

We would then obtain

$$\underline{a_n} = \Gamma_n^{-1}\gamma(n;k)$$

where

$$\Gamma_n = [\gamma_X(i-j)]_{i,j=1}^{n}$$

$$\gamma(n;k) = (\gamma_X(k), \ldots, \gamma_X(k+n-1))^T$$

$$\underline{a_n} = (a_1, \ldots, a_n)$$

This equation is called the Yule-Walker formula for prediction.

Mean Squared Error:

$$MSPE_n(k) = \gamma_X(0) - \underline{a_n}^T \gamma(n;k)$$