Time Series Data Analysis

Midterm 1

The Basics

Review

- **Random Variable** an outcome of a random phenomenon. defined according to a distribution.
- **Expectation** → population mean…
  $E[X] = \sum x p(x)$ or $= \int x p(x) dx$
- **Variance** → deviation from center
  $Var(X) = E[(X - E[X])^2]$
  $= E[X^2] - E[X]^2$
- **Joint Distribution** → depends on 2 rv's… $P(X = x, Y = y)$
  if indp… $= P(X = x) P(Y = y)$
  $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$
- **Linear Transforms**
  $aX + b$
  $E[aX + b] = aE[X] + b$
  $Var(aX + b) = a^2 Var(X)$
- **Covariance**
  $Cov(X, Y) = E[XY] - E[X]E[Y]$
  if indp then $= \emptyset$
- **Correlation**
  $Corr(X, Y) = Cov(X, Y) / \sqrt{Var X}\sqrt{Var Y}$
- **Linearity**
  $Cov(a_0 + a_1 X_1 + \cdots, b_0 + b_1 Y_1 + \cdots)$
  $= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j Cov(X_i, Y_j)$
  $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

Definitions

- **Time Series** → a set of observations x_t, each recorded at time t. discrete or continuous
- **Time Series Model** → a seq. of rv's $\{X_t\}$ that posulate $\{x_t\}$. Typically not possible to define a full probablistic model so we look at first and second order moments + properties.

models → $X_t = m_t + s_t + Y_t$ → Random Component

- **Trend component**
  > a slowly changing one direction component of a time series model
- **Seasonal component**
  > a fluctuating / periodic component of a time series model
- **Mean func, Covariance func, Autocovariance func, & Autocorrelation func** of $\{X_t\}$
  $\mu_X(t) = E[X_t]$
  $\gamma_X(r, s) = Cov(X_r, X_s)$
  $\gamma_X(h) = \gamma_X(t+h, t) = Cov(X_{t+h}, X_t)$
  $\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Cor(X_{t+h}, X_t)$
  h also called lag
  → measure how events across time are interacting
- **Weakly Stationary** → a model $\{X_t\}$ is weakly stationary if…
  i) $\mu_X(t)$ is independent of t (a constant)
  and…
  ii) $\gamma_X(t+h, t)$ is independent of t for each h. aka covariance depends on lag and not time.
- **Strict Stationary** → $(X_1, \ldots, X_n)$ and $(X_{1+h}, \ldots, X_{n+h})$ have the same joint distribution for all h and $n > 0$. strictly stationary ⇒ weakly stationary
- When we say just "stationary", we typically mean weakly stationary.

The Models

- We look at $\emptyset$ mean models first…

iid random noise

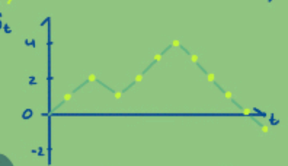
- a model w/ no trend or seasonality
- the past doesn't matter when predicting the future
- $\mu_X(t) = \emptyset$ $Var(X_t) = \sigma^2$
- $\gamma_X(h) = \begin{cases} \sigma^2, & \text{if } h = 0 \\ \emptyset, & \text{if } h \neq 0 \end{cases}$ $\rho_X(h) = \begin{cases} \emptyset, & \text{if } h \neq 0 \\ 1, & \text{if } h = 0 \end{cases}$
- iid and stationary ✓

binary process

- iid RV's $\{X_t, t = 1, 2 \ldots\}$ w/ $P(X_t = 1) = p$, $P(X_t = -1) = 1 - p$
- use this to build other processes…

random walk

- $\{S_t, t = 0, 1, 2, \ldots\}$, made by cumulativly summing iid binary rv's $\{X_t\}$. → Simple symmetric random walk, mean $\emptyset$
- ex… S_t



- $\mu_S(t) = \emptyset$ $Var(S_t) = t\sigma^2$
- $\gamma_S(h) = t\sigma^2$ ← depends on t…
- NOT stationary

white noise

- $\{X_t\}$ → uncorrelated RV's, not necessarily iid.
- $\mu_X(t) = \emptyset$ $Var(X_t) = \sigma^2$
- write as $\{X_t\} \sim WN(\emptyset, \sigma^2)$
- stationary ✓ w/ same Cov. as iid Noise.
- $\gamma_X(h) = \begin{cases} \sigma^2, & h = 0 \\ \emptyset, & h \neq 0 \end{cases}$
  $\rho_X(h) = \begin{cases} \emptyset, & h \neq 0 \\ 1, & h = 0 \end{cases}$
- every $iid(0, \sigma^2)$ is $WN(0, \sigma^2)$, but converse is not true.

first order moving average

- also written as MA(1)
- $X_t = Z_t + \theta Z_{t-1}$, w/ $\{Z_t\} \sim WN(0, \sigma^2)$
  θ is a constant
- $\mu_X(t) = \emptyset$
- $\gamma_X(h) = \begin{cases} \theta\sigma^2, & |h| = 1 \\ \sigma^2 + \theta^2\sigma^2, & h = 0 \\ \emptyset, & \text{else} \end{cases}$ $\rho_X(h) = \begin{cases} 1, & h = 0 \\ \theta / 1 + \theta^2, & |h| = 1 \\ 0, & \text{else} \end{cases}$
- stationary ✓

first order autoregression

- also written as AR(1)
- start w/ $\{X_t\}$
  $X_t = \phi X_{t-1} + Z_t$
  $\{Z_t\} \sim WN(0, \sigma^2)$ uncorrelated w/ X's
  $|\phi| < 1$, a constant
- $\mu_X(t) = \emptyset$
- $\gamma_X(h) = \frac{\phi^{|h|}\sigma^2}{1 - \phi^2}$ $\gamma_X(0) = \frac{\sigma^2}{1 - \phi^2}$
- $\rho_X(h) = \phi^{|h|}$
- stationary! ✓

classical decomposition

- how we break down our model overall…
  $X_t = m_t + s_t + Y_t$
  m_t: trend component, slowly changing
  s_t: seasonal component w/ known period d
  Y_t: random noise component, stationary
- we estimate Y_t w/ one of the stationary models described above…

The Real World

sample acf

- in real life we have data and not a predefined model.
- sample autocorrelation function → an estimate of ACF based on observable data
- $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$
- $\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x})$
- $\hat{\rho}(h) = \hat{\gamma}(h) / \hat{\gamma}(0)$

harmonics

- we could fit a parametric model to describe a seasonal component s_t.
- or sum of harmonics…
  $s_t = a_0 + \sum_{j=1}^{k}(a_j \cos(\lambda_j t) + b_j \sin(\lambda_j t))$

estimation & elimination

- we start w/ our classical decomposition model and use differencing to remove trend + seasonality.
- ∇ → lag 1-dif. operator
- $\nabla X_t = (1 - B)X_t = X_t - X_{t-1}$
  w/… $B^j(X_t) = X_{t-j}$ $\nabla^0 X_t = X_t$
  $\nabla^j X_t = \nabla(\nabla^{j-1}(X_t))$
- eliminate Trend
  > $m_t = c_0 + c_1 t + \cdots$, $\nabla m_t = \cdots = c_1$
  > any polynomial trend of degree k can be reduced to constant by ∇^k
  > $\nabla^k X_t = k! c_k + \nabla^k Y_t$
- ∇_d → lag d dif. operator
- $\nabla_d X_t = (1 - B^d)X_t = X_t - X_{t-d}$
- eliminate Seasonality
  > $\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}$

testing noise

- We need to test if our estimated noise component fits the data.
- H_0: residuals are iid.
- Sample ACF → 95% of $\hat{\rho}_i$'s should fall in $\frac{\pm 1.96}{\sqrt{n}}$
- Portmanteau Test → $Q = n\sum_{j=1}^{h}\hat{\rho}^2(j)$, $Q \sim \chi^2$
- Ljung-Box → variation of portmanteau
  $Q_{LB} = n(n+2)\sum_{j=1}^{h}\hat{\rho}^2(j)/(n-j)$ $Q_{LB} \sim \chi^2$
- McLeod-Li → variation of Ljung-Box w/ squared data
- Turning Point Test → a seq of y_i's. a turning point is when…
  T = # of turning p.
  if iid then…
  $\mu_T = 2(n-2)/3$ $\sigma_T^2 = (16n - 29)/90$
  $T \sim N(\mu_T, \sigma_T^2)$
- dif. sign test → S = # of values of $i \to y_i > y_{i-1}$
  if iid then…
  $\mu_S = \frac{1}{2}(n-1)$ $\sigma_S^2 = (n+1)/12$ $S \sim N(\mu_S, \sigma_S^2)$
- Rank Test → P = # of pairs (i, j) such that… $j > i$ and $y_j > y_i$
  if iid…
  $\mu_P = \frac{n}{4}(n-1)$
  $\sigma_P^2 = \frac{n(n-1)(2n+5)}{72}$
  $P \sim N(\mu_P, \sigma_P^2)$

good luck