Purdue University

Stationary Processes

Brockwell and Davis: Chapter 2

Forecasting

- Assume that you know X_n and need to predict X_{n+h}. This task is called **the h-step ahead forecasting**
- Terminology: n is the *forecast origin* and h is the *forecast horizon*
- Remember that, for a bivariate normal $X = (X_1, X_2) \sim N(\mu, \Sigma)$, the conditional distribution of X_1 given $X_2 = x_2$ is normal

$$N\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right) \tag{1}$$

- If the loss function involved in the definition of ”best” is the squared error loss, forecasting X_{n+h} means finding the function $m(X_n)$ such that it minimizes

$$E(X_{n+h} - m(X_n))^2$$

- We know that this is $m(X_n) = E(X_{n+h}|X_n) = \mu + \rho(h)(X_n - \mu)$ due to (1)

- The corresponding mean squared error is $E(X_{n+h} - m(X_n))^2 = \sigma^2(1 - \rho(h))^2$

- Clearly, when $|\rho(h)| \to 1$, the best linear predictor approaches $\mu \pm (X_n - \mu)$ while its mean squared error approaches 0.

Non-Gaussian case

- If we look for the best **linear** predictor only, we need to find a and b such that

$$l(X_n) = aX_n + b$$

  is minimized

- It is easy to find that

$$l(X_n) = \mu + \rho(h)(X_n - \mu)$$

Basic properties of autocovariance function

- $\gamma(0) \geq 0$
- $|\gamma(h)| \leq \gamma(0)$ for all h
- $\gamma(h) = \gamma(-h)$ (i.e. the autocovariance function is even)
- All of these are also true for autocorrelation function $\rho(\cdot)$

Main property of the autocovariance function

- A real-valued function defined on the set of all integers Z is the autocovariance function of a stationary time series iff it is even and nonnegative definite
- The function is nonnegative definite if

$$\sum_{i,j=1}^{n} a_i \kappa(i-j) a_j \geq 0$$

for all n and real-valued vectors $a = (a_1, \ldots, a_n)'$

Example

- Let us define the function

$$\kappa(h) = 1\,, h = 0$$
$$\kappa(h) = \rho\,, h = \pm 1$$
$$\kappa(h) = 0 \text{ otherwise}$$

- It is easy to realize that this is an ACVF of an MA(1) process $X_t = Z_t + \theta Z_{t-1}$ and we only need to find θ and σ^2 such that $\sigma^2(1+\theta^2) = 1$ and $\sigma^2\theta = \rho$
- It is easy to see that the real solution for θ exists iff $|\rho| \le \frac{1}{2}$

Linear Processes

- The time series X_t is a **linear process** if

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \tag{2}$$

where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $Z_t \sim WN(0, \sigma^2)$

- This is a generalization of the moving average concept; if $\psi_j \equiv 0$ for $j < 0$ we have $MA(\infty)$ instead.

- Note that the compact form of the linear process is

$$X_t = \Psi(B) Z_t$$

where $\Psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$

- Input series (white noise) $Z_t \overset{\Psi(B)}{\longrightarrow} X_t$
- For any stationary process $Y_t \overset{\Psi(B)}{\longrightarrow} X_t$ where X_t is also stationary

Why is the idea of the linear process useful

- Every stationary process can be represented as a linear process plus a deterministic component - Wold expansion; will be discussed later
- Both AR and MA processes can be represented in the form (2)
- Consider an AR(1) (autoregressive first order) process

$$X_t - \phi X_{t-1} = Z_t \tag{3}$$

where $Z_t \sim WN(0, \sigma^2)$, $|\phi| < 1$ and Z_t is uncorrelated with X_s for each $s < t$.

- It is easy to verify that the linear process

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

is

1. A solution of (3)
2. It is stationary with mean 0 and ACVF

$$\gamma_X(h) = \sum_{j=0}^{\infty} \phi^j \phi^{j+h} \sigma^2 = \frac{\sigma^2 \phi^h}{1 - \phi^2}$$

3. It is also the *only* stationary solution of (3)

Forecasting

- Consider an AR(1) model

$$r_{h+1} = \phi_0 + \phi_1 r_h + \epsilon_{h+1}$$

- The optimum point forecast of r_{h+1} is then

$$\hat{r}_h(1) = E(r_{h+1}|r_h) = \phi_0 + \phi_1 r_h$$

and the associated forecast error is

$$e_h(1) = r_{h+1} - \hat{r}_h(1) = \epsilon_{h+1}$$

- The variance of the 1-step ahead forecast error is

$$Var(\epsilon_{h+1}) = \sigma^2$$

ϵ_{h+1} is referred to as a *shock* to the series r_h at a time $h+1$

- For a more general AR(p) model, we have

$$r_{h+1} = \phi_0 + \sum_{i=1}^{p} \phi_i r_{h-i} + \epsilon_{h+1}$$

with the optimal forecast

$$\hat{r}_h(1) = E(r_{h+1}|F_h) = \phi_0 + \sum_{i=1}^{p} \phi_i r_{h-i}$$

- The forecast error is again ϵ_{h+1} and its variance is σ^2.

- Consider an AR(p) model

$$r_{h+2} = \phi_0 + \phi_1 r_{h+1} + \ldots + \phi_p r_{h+2-p} + \epsilon_{h+2}$$

where the 2-step ahead forecast is

$$\hat{r}_h(2) = E(r_{h+2}|F_h) = \phi_0 + \phi_1 \hat{r}_h(1) + \phi_2 r_h + \ldots + \phi_p r_{h+2-p}$$

- Its error is

$$e_h(2) = r_{h+2} - \hat{r}_h(2) = \epsilon_{h+2} + \phi_1 \epsilon_{h+1}$$

with the variance $Var\, e_h(2) = (1 + \phi_1^2)\sigma^2 \geq Var\, e_h(1)$

- The forecast precision decreases as the forecast horizon increases!

- In general, we have

$$r_{h+l} = \phi_0 + \phi_1 r_{h+l-1} + \ldots + \phi_p r_{h+l-p} + \epsilon_t$$

- The l-step ahead forecast is

$$\hat{r}_h(l) = \phi_0 + \sum_{i=1}^{n} \phi_i \hat{r}_h(l-i)$$

- It can be shown that for a stationary AR(p) model, $\hat{r}_h(l)$ converges to $E(r_t)$ as $l \to \infty$. This is called the **mean reversion** property in the finance literature. For an AR(1) model, the speed of mean reversion is measured by the **half-life** defined as $k = \log(\frac{0.5}{|\phi_1|})$; the variance of the forecast error then approaches the unconditional variance of r_t.