1

Time Series Analysis:-

Forecasting Accuracy:-

i) Forecast Errors.
ii) Diebold Mariano Test.
iii) Pesaran Timmerman Test

Forecast Error:- If $y_1, y_2, \ldots, y_n$ represents a time series then $\hat{y}_i$ represents the i^{th} forecasted value where $i \le n$.
For $i \le n$ the i^{th} error e_i is then:

$$e_i = y_i - \hat{y}_i$$

Goal is to find a forecast that minimize the errors.

A number of ~~errors~~ measures are commonly used to determine the accuracy of a forecast, including the mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i|$$ → Mean Absolute Error.
→ Mean Absolute Deviation (MAD)

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2$$ → Mean Squared Error

$$RMSE = \sqrt{MSE}$$ → Root Mean Squared Error.
In other words we can say Standard Deviation is the Root Mean Squared Deviation.

Other Measurements:

Mean Absolute ~~Percentage~~ Scaled Error:

$$MASE = 1/n \sum_{i=1}^{n}|e_i| \Big/ \frac{1}{n-1}\sum_{i=2}^{n}|y_i - y_{i-1}|$$

Mean Absolute Percentage Error.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{e_i}{y_i}\right|$$

Symmetric Mean Absolute Percentage Error:

$$SMAPE = 1/n \sum_{i=1}^{n} |e_i| / (|y_i| + |\hat{y}_i|)/2$$

(2) Diebold - Mariano Test —

Let us suppose that we have two forecasts: $f_1, f_2, \ldots, f_n$ and $g_1, g_2, \ldots, g_n$ for the time series $y_1, y_2, \ldots, y_n$. Objective is to find better forecast having better predictive accuracy, i.e. to find smaller error measurement.

Diebold-Mariano Test to find out whether the forecasts are significantly different.

Let us consider e_i and r_i be the residuals for the two forecasts, i.e.

$$e_i = y_i - f_i \qquad r_i = y_i - g_i$$

d_i (loss differential) is defined as one of the following

$$d_i = e_i^2 - r_i^2 \quad \text{or} \quad d_i = |e_i| - |r_i|$$

MSE statistics   MAE statistics

We now define $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i \qquad \mu = E[d_i]$

For $n > k \geq 1$, define

$$\gamma_k = \frac{1}{n}\sum_{i=k+1}^{n}(d_i - \bar{d})(d_{i-k} - \bar{d})$$

where γ_k is the auto covariance at lag k.

(*) The auto correlation function (ACF) at lag k is denoted by ρ_k of a stationary stochastic process which is defined as

Def $\rightarrow$ $\rho_k = \frac{\gamma_k}{\gamma_0}$ where $\gamma_k = \operatorname{cov}(y_i, y_{i+k})$ for any i.

γ_0 is the variance of the stochastic process.

(3)

Defn 2 → The mean of a time series: $y_1, y_2, \cdots, y_n$ is

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

The AutoCovariance function at lag k, for $k \geq 0$ of the time series is defined by:

$$S_k = \frac{1}{n}\sum_{i=1}^{n-k}(y_i - \bar{y})(y_{i+k} - \bar{y}) = \frac{1}{n}\sum_{i=k+1}^{n}(y_i - \bar{y})(y_{i-k} - \bar{y})$$

The Auto Correlation Function (ACF) at lag k for $k \geq 0$ of the time series is defined by:

$$r_k = \frac{S_k}{S_0}$$

---

for $h \geq 1$ the Diebold-Mariano ~~Test~~ Statistics:

$$DM = \frac{\bar{d}}{\sqrt{\left[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k\right]/n}}$$

It is generally sufficient to use the value $h = n^{1/3} + 1$

Under the assumption that $\mu = 0$ (the null hypothesis) DM follows the Standard Normal Distribution:

$$DM \sim N(0,1)$$

Thus there is a significant difference between the forecasts if $|DM| > Z_{crit}$ where Z_{crit} is the two tailed critical value for the standard normal distribution.

$$Z_{crit} = \text{NORM.S.DIST}(1 - \alpha/2, \text{TRUE})$$

Key assumption for Diebold-Mariano test is that the loss differential time series d_i is stationary.

Diebold-Mariano test tends to reject the null hypothesis too often for a small sample. A better test is Harvey-Leybourne-Newbold (HLN) test.

$$HLN = DM\sqrt{[n+1-2h+h(h-1)]/n} \sim T(n-1)$$

(1) Pesaran-Timmermann Test

The Pesaran-Timmermann Test is to determine whether a forecast does a good job in predicting the change in direction of a time series, i.e. the directional accuracy.

For any time series t_i with n elements we first define

$$d_t(i) = \begin{cases} 1; & t_i > 0 \\ 0; & t_i \leq 0 \end{cases} \quad p_t = \frac{1}{n}\sum_{i=1}^{n} d_t(i) \quad q_t = \frac{p_t(1-p_t)}{n}$$

Let us suppose we have a time series y_i with n elements which is forecasted by z_i and define

$$p = p_y p_z + (1-p_y)(1-p_z)$$

$$v = \frac{p(1-p)}{n} \quad w = (2p_y - 1)^2 q_z + (2p_z - 1)^2 q_y + 4q_y q_z$$

Under the null hypothesis that z does not forecasts the direction of the change in y (i.e. the sign of y_i) we have the following test statistics:

$$PT = \frac{p_{yz} - p}{\sqrt{v-w}} \sim N(0,1)$$

The ~~Pesa~~ Pesaran-Timmermann test is a one-tailed test in which the critical region (where the null hypothesis is rejected) is the upper tail of the standard normal distribution. Thus if $1 -$ NORM.S.DIST(PT, TRUE) $< \alpha$ then we can reject the Null Hypothesis and state that $1-\alpha$ confidence that the forecast accurately predicts the sign y_i.

If the sign of ~~all~~ all the elements in y_i (or z_i) are the same, then PT will be undefined.

(5) ## Simple Moving Average Model:

We forecast the next value (s) in a time series based on the average of a fixed finite number m of the previous values. Thus ~~for~~ $\forall\ i > m$

$$\hat{y}_i = \frac{1}{m} \sum_{j=i-m}^{i-1} y_i = (y_{i-m} + \cdots + y_{i-1})/m$$

Weighted Moving Average Model:

In simple Moving Average Forecast the weights given to the previous three values were all equal. Now we are considering the case where these weights can be different. This type of forecasting is called weighted Moving Average. Here we assign m weights $w_1, w_2, \ldots w_m$ where $w_1 + w_2 + \cdots + w_m = 1$ and defining the forecasted value as:

$$\hat{y}_i = w_m y_{i-m} + \cdots + w_1 y_{i-1}$$

In the simple moving average all the weights are equal to $1/m$.

Simple Exponential Smoothing:-

Exponential Smoothing improves on weighted Moving Average by taking all previous observations into account, while still favoring the most recent observation.

Simple/Single Exponential Smoothing, the forecasted value at time $i+1$ is based on the value at time i and the forecasted value at time i (and so indirectly on all the previous time value). In particular, for some α where $0 \leq \alpha \leq 1$, $\forall\ i > 1$, we define

$$\hat{y}_1 = y_1 \quad \hat{y}_{i+1} = \hat{y}_i + \alpha e_i$$

⑥ ~~The~~ $\hat{y}_1 = y_1$ and $\hat{y}_{i+1} = \hat{y}_i + \alpha e_i$

the above iteration can also be expressed in algebra as:

$$\hat{y}_1 = y_1 \quad \hat{y}_{i+1} = \alpha y_i - \alpha y_i + \hat{y}_i$$

by this $\hat{y}_1 = y_1$, $\alpha y_i = \alpha \hat{y}_i$

$$= \alpha y_i - \alpha \hat{y}_i + \hat{y}_i$$

$$= \alpha y_i + (1-\alpha)\hat{y}_i$$

$$\therefore \hat{y}_5 = \alpha y_4 + (1-\alpha)\hat{y}_4 = \alpha y_4 + (1-\alpha)[\alpha y_3 + (1-\alpha)\hat{y}_3]$$

$$= \alpha y_4 + \alpha(1-\alpha) y_3 + (1-\alpha)^2 \hat{y}_3$$

$$= \alpha y_4 + \alpha(1-\alpha) y_3 + (1-\alpha)^2 [\alpha y_2 + (1-\alpha)\hat{y}_2]$$

$$= \alpha y_4 + \alpha(1-\alpha) y_3 + \alpha(1-\alpha)^2 y_2 + (1-\alpha)^3 \hat{y}_2$$

$$= \alpha y_4 + \alpha(1-\alpha) y_3 + \alpha(1-\alpha)^2 y_2 + (1-\alpha)^3 [\alpha y_1 + (1-\alpha)\hat{y}_1]$$

$$= \alpha y_4 + \alpha(1-\alpha) y_3 + \alpha(1-\alpha)^2 y_2 + \alpha(1-\alpha)^3 y_1 + (1-\alpha)^4 \hat{y}_1$$

~~Holt's Linear Trend:~~

(7) Holt's Linear Trend:– A distinct upward/downward trend can't be captured using Single Exponential Smoothing. To capture that we need Holt's Linear Trend Method a.k.a Double Exponential smoothing. This is accomplished by adding a second single Exponential Smoothing model to capture the trend.

Let us consider: $u_1 = y_1 \quad v_1 = 0$

$$u_i = \alpha y_i + (1-\alpha)(u_{i-1} + v_{i-1})$$
$$v_i = \beta(u_i - u_{i-1}) + (1-\beta)v_{i-1}$$
$$\hat{y}_{i+1} = u_i + v_i \quad \text{where } 0 < \alpha \leq 1 \quad 0 \leq \beta \leq 1$$

The alternative form of equation:

$$u_1 = y_1 \qquad v_1 = 0$$
$$u_i = u_{i-1} + v_{i-1} + \alpha e_i$$
$$v_i = v_{i-1} + \alpha\beta e_i$$
$$\hat{y}_{i+1} = u_i + v_i \quad \text{where } e_i = y_i - (u_{i-1} + v_{i-1}) = y_i - \hat{y}_{i-1}$$

If $\beta = 0$, Holt's model is equivalent to Single Exponential Smoothing

Holt's Winter Method:– This is also known as Triple Exponential Smoothing. Here a seasonal component is added to the Holt's Linear Trend Model (Double exponential Smoothing).

Let c be the length of the seasonal cycle. Thus $c = 12$ for months in a year, $c = 7$ for days in a week, $c = 4$ for quarters in a year.

$$u_i = \alpha(y_i/s_{i-c}) + (1-\alpha)(u_{i-1} + v_{i-1}) \qquad u_c = y_c/s_c \quad v_c = 0$$
$$v_i = \beta(u_i - u_{i-1}) + (1-\beta)v_{i-1} \qquad s_i = \frac{y_i}{\frac{1}{c}\sum_{j=1}^{c} y_j}$$
$$s_i = \gamma(y_i/u_i) + (1-\gamma)s_{i-c}$$
$$\hat{y}_{i+1} = (u_i + v_i)s_{i+1-c}$$

where $0 < \alpha \leq 1, 0 \leq \beta \leq 1$ and $0 \leq \gamma \leq 1-\alpha$. The initial values $1 \leq i \leq c$ are given

(8) The alternative form of the equation is given by:

$$u_i = u_{i-1} + v_{i-1} + \alpha e_i$$

$$v_i = v_{i-1} + \alpha\beta e_i$$

$$s_i = s_{i-c} + \gamma e_i$$

$$\hat{y}_{i+1} = u_i + v_i$$

where $e_i = y_i - (u_{i-1} + v_{i-1} + s_{i-c}) = y_i - \hat{y}_{i-1}$

If $\gamma = 0$, then Holt-Winter's model is equivalent to Holt's Linear Trend Model and if $\beta = 0$ and $\gamma = 0$ then Holt-Winters model is equivalent to Single Exponential Smoothing model.

Scanned by CamScanner

(9) Stationary Process –

A stochastic process (random process) is stationary if the properties of the time series (i.e. mean, variance, ~~autocovariance etc.) are the same~~ autocovariance etc are all constant) when measured from any two starting points in time are constant.

Mean: $E[y_i] = \mu$,

Variance: $var(y_i) = E[(y_i - \mu)^2] = \sigma^2$

Auto covariance for any lag k: $cov(y_i, y_{i+k}) = E[(y_i - \mu)(y_{i+k} - \mu)] = \gamma_k$

Auto Correlation Function at lag k : It is denoted by ρ_k of a stationary stochastic process which is defined as:

$\rho_k = \gamma_k / \gamma_0$ where $\gamma_k = cov(y_i, y_{i+k})$ for any i.

γ_0 = variance of the stochastic process.

Mean: The mean of time series $y_1, y_2, \ldots, y_n$ is

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

The Auto Covariance Function at lag k for $k \geq 0$ of the time series is defined by:

$$s_k = \frac{1}{n} \sum_{i=1}^{n-k} (y_i - \bar{y})(y_{i+k} - \bar{y}) = \frac{1}{n} \sum_{i=k+1}^{n} (y_i - \bar{y})(y_{i-k} - \bar{y})$$

$$r_k = \frac{s_k}{s_0}$$

(10) Partial Auto Correlation Function.

For regression of y on x_1, x_2, x_3, x_4 the partial correlation between y and x_1 is

$$\frac{cov(y, x_1 | x_2, x_3, x_4)}{\sqrt{var(y|x_2, x_3, x_4) \cdot var(x_1|x_2, x_3, x_4)}}$$

This can be calculated as the correlation between the residuals of the regression of y on x_2, x_3, x_4 with residuals of x_1 on x_2, x_3, x_4.

For the time series, the h^{th} order partial auto correlation is the partial correlation of y_i with y_{i-h} conditional on $y_{i-1}, y_{i-2}, \dots, y_{i-h+1}$ i.e

$$\frac{cov(y_i, y_{i-h} | y_{i-1}, \dots, y_{i-h+1})}{\sqrt{var(y_i | y_{i-1}, \dots y_{i-h+1}) \& var(y_{i-h} | y_{i-1} \dots y_{i-h+1})}}$$

For $k>0$, the Partial Auto Correlation function (PACF) of order k is denoted by π_k of a stochastic process and is defined as the k^{th} element in the column vector $\Gamma_k^{-1} \delta_k$. Here Γ_k is the $k \times k$ auto covariance matrix $\Gamma_k = [v_{ij}]$ where $v_{ij} = \gamma_{|i-j|}$ and δ_k is the $k \times 1$ column vector $\delta_k = [\gamma_i]$. We also define $\pi_0 = 1$. We can also define π_{ki} to be the ith element in the vector $\Gamma_k^{-1} \delta_k$ and so $\pi_k = \pi_{kk}$

(11) Provided $\gamma_0 > 0$ the partial correlation function of order k is equal to k^{th} element in the following column matrix divided by γ_0

$$\Sigma_k^{-1} \tau_k$$

Here Σ_k is the k×k autocorrelation matrix $\Sigma_k = [\omega_{ij}]$ where $\omega_{ij} = \rho_{|i-j|}$ and τ_k is the k×1 column vector $\tau_k = [\rho_i]$

If $\gamma_0 > 0$ then Σ_k and Γ_k are invertible for all m.

The Partial Auto Correlation Function (PACF) or order k is denoted by p_k of a time series, is defined in a similar manner as the last element in the following matrix divided by r_0

$$R_k^{-1} C_k$$

Here R_k is the k×k matrix $R_k = [s_{ij}]$ where $s_{ij} = r_{|i-j|}$ and C_k is the k×1 column vector $C_k = [r_i]$

We also define $p_0 = 1$ and p_{ik} to be the ith element in the matrix $R_k^{-1} C_k$ and so $p_k = p_{kk}$. These values can also be calculated from the autocovariance matrix of the time series in the same manner as decribed above

(*) The first difference $z_i = y_i - y_{i-1}$ of a random walk is stationary since it takes form

$$z_i = y_i - y_{i-1} = y_0 + \delta i + \sum_{j=1}^{i} \varepsilon_j - y_0 - \delta(i-1) - \sum_{j=1}^{i-1} \varepsilon_j = \delta + \epsilon_i$$

⇒ Purely Random Time Series (White Noise)

(11)

Purely Random Time Series (White Noise)

A purely random time series $y_1, y_2, \cdots, y_n$ (a.k.a white noise) takes form

$$y_i = \mu + \varepsilon_i \quad \text{where} \quad \varepsilon_i \sim N(0,\sigma)$$
$$cov(\varepsilon_i, \varepsilon_j) = 0 \quad \text{for } i \neq j$$

Clearly $E[y_i] = \mu$, $var(y_i) = \sigma^2$ and $cov(y_i, y_j) = 0$ for $i \neq j$. Since these values are constants, this type of time series is stationary. Also note that $\rho_h = 0 \ \forall h > 0$.

Random Walk: A random walk time series $y_1, y_2, \cdots y_n$ takes the form $y_i = \delta + y_{i-1} + \varepsilon_i$ $\quad \varepsilon_i \sim N(0,\sigma)$
$cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$.

If $\delta = 0$, then the random walk is said to be without drift, while $\delta \neq 0$, then the random walk is with drift (i.e. the drift equals to δ).

for $i > 0$ $\quad y_i = y_0 + \delta i + \sum_{j=1}^{i} \varepsilon_j$

$E[y_i] = y_0 + \delta i$, $var(y_i) = \sigma^2 i$ and $cov(y_i, y_j) = 0$ for $i \neq j$.
The variance value is not constant but vary with time i, so this type of time series is not stationary. The mean values are constant only for a random walk without drift.

$cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$

$$cov(y_i, y_{i-h}) = E\left[\sum_{j=1}^{i} \varepsilon_j \cdot \sum_{j=1}^{i-h} \varepsilon_j\right] = E\left[\sum_{j=1}^{i-h} \varepsilon_j^2\right] = (i-h)\sigma^2$$

$$corr(y_i, y_{i-h}) = \frac{cov(y_i, y_{i-h})}{var(y_i)\, var(y_{i-h})} = \frac{(i-h)\sigma^2}{\sqrt{\sigma^2 i}\sqrt{\sigma^2 (i-h)}} = \frac{i-h}{\sqrt{i(i-h)}} = \sqrt{\frac{i-h}{i}}$$

(*) Pg 11

(13) Deterministic Trend.

A time series with linear deterministic trend can be modeled as

$$y_i = \mu + \delta i + \varepsilon_i$$

Now $E[y_i] = \mu + \delta i$ and $var(y_i) = \sigma^2$, variance is constant but mean varies with time i. This is also not stationary.

This type of time series can be transformed into stationary time series by detrending, i.e. by setting $z_i = y_i - \delta i$. In this case $z_i = \mu + \varepsilon_i$ which is a purely random time series (White Noise).

~~Quadratic Deterministic Trend: $y_i = \mu +$~~

The type of Random walks described previously are said to have a stochastic trend. We can also have a random walk with deterministic trend.

$$y_i = \mu + y_{i-1} + \delta i + \varepsilon_i \rightarrow \delta \text{ is constant.}$$

(14) Dickey-Fuller Test —

We consider the stochastic process of form

$$y_i = \phi y_{i-1} + \varepsilon_i \quad \text{where } |\phi| \le 1$$

and ε_i is White Noise (Purely Random Time Series).

If $|\phi| = 1$, we have it called a unit root.

In particular if $\phi = 1$, we have the random walk (without drift)

which is not stationary.

If $|\phi| = 1$ the process is not stationary.

If $|\phi| < 1$ the process is stationary.

If $|\phi| > 1$ the process is called explosive and increases over time.

The Dickey-Fuller test is a way to determine whether the above process has unit root.

First difference $\quad y_i - y_{i-1} = \phi y_{i-1} + \varepsilon_i - y_{i-1}$

$$= (\phi - 1) y_{i-1} + \varepsilon_i$$

If we use delta operator, defined by $\Delta y_i = y_i - y_{i-1}$ and set $\beta = \phi - 1$ then the equation becomes linear regression equation.

$$\Delta y_i = \beta y_{i-1} + \varepsilon_i$$

Moving Average process

A q. order moving average process, denoted by MA(q) takes the form:

$$y_i = \mu + \varepsilon_i + \theta_1 \varepsilon_{i-1} + \cdots + \theta_q \varepsilon_{i-q}$$

The subscript i is considered the representation of time.

It can be observed that the value of y at time $i+1$ is a linear function of past errors. We assume error terms are independently distributed with a normal distribution with mean zero and constant variance σ^2. Thus:

$$\varepsilon_i \sim N(0, \sigma) \qquad cov(\varepsilon_i, \varepsilon_j) = 0 \text{ if } i \neq j.$$

An MA(q) process can be expressed as: $z_i = \varepsilon_i + \theta_1 \varepsilon_{i-1} + \cdots + \theta_q \varepsilon_{i-q}$

where $z_i = y_i - \mu$

Simplifying the analysis by restricting for case of mean $= 0$. So by using the lag operator, we can express a zero mean MA(q) process as:

$$z_i = \theta(L)\,\varepsilon_i$$

$$\text{where } \theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

→ The mean of a MA(q) process is μ.

→ The variance of a MA(q) process is

$$\cancel{var(y_i)}$$
$$var(y_i) = \sigma^2 (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

→ The Auto Correlation Function (ACF) of an MA(1) process:

$$\rho_1 = \frac{\theta_1}{1+\theta_1^2} \qquad \rho_h = 0 \text{ for } h > 1.$$

(16)

→ The ~~Auto Corre~~ Auto Correlation Function (ACF) of an MA(2) process is,

$$\rho_1 = \frac{\theta_1 + \theta_1\theta_2}{1+\theta_1^2+\theta_2^2} \qquad \rho_2 = \frac{\theta_2}{1+\theta_1^2+\theta_2^2} \qquad \rho_h = 0 \text{ for } h>2$$

→ The AutoCorrelation Function (ACF) of an MA(q) process is

$$\rho_h = \frac{\theta_h + \sum_{j=1}^{q-h} \theta_j \theta_{j+h}}{1+\sum_{j=1}^{q} \theta_j^2} \quad \text{for } h \le q \text{ and } \rho_h = 0 \text{ for } h>q$$

→ The PACF of an MA(1) process:

$$\pi_K = \frac{-(-\theta_1)^k}{1+\sum_{i=1}^{k} \theta_1^{2i}} \qquad \pi_{k,k-j} = \frac{-(-\theta_1)^j}{1+\sum_{i=1}^{j} \theta_1^{2i}} \cdot \pi_K$$

where $1 \le j < n$.

If the process is invertible:

$$\pi_K = \frac{-(-\theta_1)^k(1-\theta_1^2)}{1-\theta_1^{2(i+1)}}$$

Infinite order Moving Average process, denoted by MA(∞) takes the form:

$$y_i = \mu + \varepsilon_i + \sum_{j=1}^{\infty} \psi\, \varepsilon_{i-j}$$

where the following infinite series is finite (i.e. converges to a real value

$$\sum_{j=1}^{\infty} |\psi_j| < \infty \quad \text{and } \varepsilon_i \sim N(0,\sigma)$$

$$\mathrm{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ for } i \ne j.$$

We can express ~~M(~~ MA(∞) process as:

$$y_i = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{i-j} \qquad \text{It is assumed that } \psi_0 = 1$$

Observation: $\sum_{j=0}^{\infty} |\psi_j|$ converges, which ensures that the y_i takes finite values and $\sum_{j=0}^{\infty} \psi_j^2$ converges.

Scanned by CamScanner

(17) Any stationary AR(1) process canbe expressed as MA(∞) process

$$\psi_j = \phi_1^j$$

for AR(1) Process $y_i = \phi_0 + \phi_1 y_{i-1} + \varepsilon_i$ canbe expressed as

$$y_i = \mu + \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{i-j}$$

where $\mu = \dfrac{\phi_0}{1-\phi_1}$

<u>Another approach</u> — Using the lag operator.

An AR(1) process with 0 mean can be expressed as:

$\phi(L) y_i = \varepsilon_i$ where $\phi(L) = \underline{\underline{1 - \phi_1 L}}$

and $y_i = \psi(L) \varepsilon_i$

where $\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$

Substituting the first equation in the 2nd

$$y_i = \psi(L) \varepsilon_i = \psi(L) \phi(L) y_i$$

i.e. $$1 = \psi(L)\phi(L) = (1 - \phi L) \sum_{j=0}^{\infty} \psi_j L^j = \sum_{j=0}^{\infty} \psi L^j - \sum_{j=0}^{\infty} \phi_1 \psi_j L^{j+1}$$

$$= 1 + \sum_{j=1}^{\infty} (\psi_j - \phi_1 \psi_{j-1}) L^j$$

We know $\psi_0 = 1$ Equating the coefficients, we see that $\forall j > 0$

$$\cancel{\psi_j} \quad \psi_j - \phi_1 \psi_{j-1} = 0$$

Thus $$\psi_1 = \phi_1 \psi_0 = \phi_1$$

$$\psi_2 = \phi_1 \psi_1 = \phi_1 \phi_1 = \phi_1^2 \quad \therefore \psi_n = \phi_1^n$$

So $$y_i = \psi(L) \epsilon_i = \sum_{j=0}^{\infty} \psi_j L^j \epsilon_j = \sum_{j=0}^{\infty} \phi_1^j L^j \epsilon_j$$

we know $1 = \psi(L)\phi(L)$

$$\therefore \psi(L) = \frac{1}{\phi(L)}$$

(18) <u>Wold's Decomposition Theorem.</u>

Any Stationary process can be represented as MA(∞) process.

→ The following are true for MA(∞) process

$$E[y_i] = \mu$$

$$\mathrm{var}(y_i) = \gamma_0 = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 \geq 0$$

$$\gamma_k = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

$$\rho_k = \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \Big/ \sum_{j=0}^{\infty} \psi_j^2 \Rightarrow \text{Correlation coeffi}$$

<u>Invertibility of MA(q) processes</u> —

Like Infinite Order Moving Average MA(∞) process can infinite order Auto Regressive AR(∞) process can be defined.

Let us suppose we have MA(1) process with $\mu = 0$.

$$y_i = \varepsilon_i + \theta_1 \varepsilon_{i-1}$$

$$\varepsilon_i = y_i - \theta_1 \varepsilon_{i-1} = y_i - \theta_1 (y_{i-1} - \theta_1 \varepsilon_{i-2})$$

$$= y_i - \theta_1 y_{i-1} + \theta_1^2 \varepsilon_{i-2}$$

Continuing in this way, after n steps,

$$\varepsilon_i = y_i - \theta_1 y_{i-1} + \theta_1^2 y_{i-2} - \theta_1^3 y_{i-3} + \cdots + (-\theta_1)^n y_{i-n} + (-\theta_1)^{n+1} \varepsilon_{i-n-1}$$

$$= \sum_{j=0}^{n} (-\theta_1)^j y_{i-j} + (-\theta_1)^{n+1} \varepsilon_{i-n-1}$$

So we have

$$\varepsilon_i = y_i - \theta_1 y_{i-1} + \theta_1^2 y_{i-2} - \theta_1^3 y_{i-3} + \cdots = \sum_{j=0}^{\infty} (-\theta_1)^j y_{i-j}$$

or equivalently,

$$y_i = \theta_1 y_{i-1} - \theta_1^2 y_{i-2} + \cdots + \varepsilon_i = -\sum_{j=1}^{\infty} (-\theta_1)^j y_{i-j} + \varepsilon_i$$

(19)

~~Calculating MA coefficients using ACF~~

If the process contains both AutoRegressive: AR(p) and Moving Average: MA(q), then the process is called AutoRegressive Moving Average: ARMA(p, q) and can be ~~explained as~~: expressed as:

$$y_i = \phi_0 + \phi_1 y_{i-1} + \phi_2 y_{i-2} + \dots + \phi_p y_{i-p} + \varepsilon_i + \theta_1 \varepsilon_{i-1} + \dots + \theta_q \varepsilon_{i-q}$$

$$\therefore y_i = \phi_0 + \sum_{j=1}^{p} \phi_j y_{i-j} + \varepsilon_i + \sum_{j=1}^{q} \theta_j \varepsilon_{i-j}$$

We can define ARMA(p, q) process with zero mean by removing the constant term (i.e. ϕ_0) and saying that $y_1, y_2, \dots, y_n$ has an ARMA(p, q) with mean μ if the time series $z_1, z_2, \dots, z_n$ has an ARMA(p, q) process with 0 mean where $z_i = y_i - \mu$.

If we include the constant term then as in AR(p) case, for a stationary ARMA(p, q) process

$$\mu = \frac{\phi_0}{1 - \sum_{j=1}^{p} \phi_j}$$

An equivalent expression for an ARMA(p, q) process with zero mean is $y_i - \sum_{j=1}^{p} \phi_j y_{i-j} = \varepsilon_i + \sum_{j=1}^{q} \theta_j \varepsilon_{i-j}$

which can be expressed as the lag or back shift operator

$$\phi(L) y_i = \Theta(L) \varepsilon_i$$

$$y_i = \frac{\Theta(L)}{\phi(L)} \epsilon_i$$

Scanned by CamScanner

(20)

ARMA(1,1) process.

$$\phi(L) = 1 - \phi_1 L \qquad \theta(L) = \cancel{1+\theta(L)}\ 1 + \theta_1 L$$

Let us suppose that $|\phi_1| < 1$. To create ~~a~~ MA(∞) process

$$y_i = \frac{\theta(L)}{\phi(L)} \varepsilon_i = \frac{1+\theta_1 L}{1 - \phi_1 L} \varepsilon_i = (1+\theta_1 L)(1 + \phi_1 L + \phi_1^2 L^2 + \cdots).$$

$$= 1 + (\theta_1 + \phi_1) L + (\theta_1 + \phi_1)\phi_1 L^2 + (\theta_1 + \phi_1)\phi_1^2 L^3 + \cdots$$

$$y_i = \left(\sum_{j=0}^{\infty} \psi_j L^j\right) \varepsilon_i \quad \text{where } \psi_0 = 1 \text{ and for } j > 0.$$

$$\psi_j = (\theta_1 + \phi_1)\phi_1^{j-1}$$

So MA(∞) process is $\Rightarrow$ $y_i = \varepsilon_i + (\theta_1 + \phi_1) \sum_{j=1}^{\infty} \phi_1^{j-1} \varepsilon_{i-j}$