MAFS5130 Time Series Analysis

Thompson Hu

1 Basic concepts

The foundation of time series analysis is stationarity.

1.1 Strictly stationary process

For n and $t_1, t_2, \cdots, t_n$, if

$$\left(X_{t_1}, \cdots, X_{t_n}\right) \stackrel{dist}{=} \left(X_{t_1+k}, \cdots, X_{t_n+k}\right), \forall k$$

we say $\{X_t\}$ is strictly stationary process.

1.2 Weakly stationary process

If $\{X_t\}$ satisfied that

- $E\left(X_t^2\right) < \infty$;
- $E(X_t) = \mu$, which is constant;
- $\operatorname{Cov}\left(X_t, X_{t+k}\right)$ is not related to t for k.

we say X_t is weakly stationary process.

1.3 Autocorrelation function (ACF)

Supposed X_t is weakly stationary sequence, then

$$\rho\left(X_{t-k}, X_t\right) = \frac{\operatorname{Cov}\left(X_{t-k}, X_t\right)}{\sqrt{\operatorname{Var}\left(X_{t-k}\right)\operatorname{Var}\left(X_t\right)}} = \frac{\gamma_k}{\sqrt{\gamma_0\gamma_0}} = \frac{\gamma_k}{\gamma_0}, k = 0, 1, \ldots, \forall t \tag{1}$$

We say $\{\rho_k, k = 0, 1, \ldots\}$ is autocorrelation function (ACF) of $\{X_t\}$ and $\rho_0 = 1$.

If weakly stationary sequence $\{X_t\}$ satisfy that $\rho_k = 0, k = 1, 2, \ldots$, then $\{X_t\}$ is **white noise** sequence.

From the definition (1), we know

$$\rho_k = \frac{\operatorname{Cov}\left(X_{t-k}, X_t\right)}{\operatorname{Var}\left(X_t\right)} = \frac{\operatorname{Cov}\left(X_t, X_{t-k}\right)}{\operatorname{Var}\left(X_t\right)} = \rho_{-k}$$

1.4 Partial ACF (PACF)

Supposed X_t is a stationary sequence, the conditional correlation

$$\operatorname{Corr}\left(X_t, X_{t+k} | X_{t+1}, \cdots, X_{t+k-1}\right)=\frac{\operatorname{Cov}\left[\left(X_t-\hat{X}_t\right)\left(X_{t+k}-\hat{X}_{t+k}\right)\right]}{\sqrt{\operatorname{Var}\left(X_t-\hat{X}_t\right) \operatorname{Var}\left(X_{t+k}-\hat{X}_{t+k}\right)}}$$

is called the PACF of X_t and X_{t+k}, denoted by ϕ_{kk}, where $\hat{X}_t=E\left(X_t | X_{t+1}, \cdots, X_{t+k-1}\right)$.

From the formula : $\phi_{11}=\rho_1$ and

$$\phi_{kk}=\frac{\left|\begin{array}{cccccc} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ & & \cdots & & & \\ & & \cdots & & & \\ & & \cdots & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{array}\right|}{\left|\begin{array}{cccccc} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ & & \cdots & & & \\ & & \cdots & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{array}\right|}$$

we know the key feature is that PACF cuts off at lag p for an AR(p) model.

1.5 Ljung-Box test

Box and Pierce presented portmanteau statistic

$$Q_*(m)=T \sum_{j=1}^m \hat{\rho}_j^2 \tag{2}$$

in 1970 to test

$$H_0: \rho_1=\cdots=\rho_m=0 \longleftrightarrow H_a: \rho_i \neq 0 \text{ for some } i$$

In 1978, Ljung and Box improved such method by changing previous statistic (2) as

$$Q(m)=T(T+2) \sum_{j=1}^m \frac{\hat{\rho}_j^2}{T-j}$$

If $Q(m)>\chi_m^2(\alpha)$ or p-value is less than α, then we reject H_0.

1.6 Linear time series

X_t is linear if X_t can be written as

$$X_t=\mu+\sum_{i=0}^{\infty} \psi_i Z_{t-i}$$

where μ is a constant, $\psi_0=1$ and $\{Z_t\}$ is a sequence of white noises.

1.7 Information criterion

- Akaike information criterion

  $$\operatorname{AIC}(k)=\ln\left(\hat{\sigma}_k^2\right)+\frac{2k}{T}$$

  for an AR(k) model, where σ_k^2 is the MLE of residual variance.

- BIC criterion

  $$\operatorname{BIC}(k)=\ln\left(\hat{\sigma}_k^2\right)+\frac{k \ln(T)}{T}$$

2 ARMA model

2.1 Auto Regression (AR) model

2.1.1 AR(1) model

The form of AR(1) can be written as:

$$X_t = \phi_0 + \phi_1 X_{t-1} + Z_t \tag{3}$$

where ϕ_0 and ϕ_1 are real numbers, which are referred to as parameters. Z_t is white noise with mean zero and variance σ^2.

Conditional expectation:

$$E\left(X_t | X_{t-1}\right) = E\left(\phi_0 + \phi_1 X_{t-1} + Z_t | X_{t-1}\right) = \phi_0 + \phi_1 X_{t-1}$$

Under the stationarity[1] condition, we can get

$$\mu = \phi_0 + \phi_1 \mu \longrightarrow E\left(X_t\right) = \mu = \frac{\phi_0}{1-\phi_1}$$

Conditional variance:

$$\operatorname{Var}\left(X_t | X_{t-1}\right) = \sigma^2$$

Under the stationarity condition, we can get

$$\begin{aligned} \operatorname{Var}\left(X_t\right) &= \operatorname{Var}\left(\phi_0 + \phi_1 X_{t-1} + Z_t\right) \\ &= \phi_1^2 \operatorname{Var}\left(X_{t-1}\right) + \operatorname{Var}\left(Z_t\right) \\ &= \phi_1^2 \operatorname{Var}\left(X_{t-1}\right) + \sigma^2 \end{aligned}$$

so we finally get[2]

$$\operatorname{Var}\left(X_t\right) = \frac{\sigma^2}{1-\phi_1^2}$$

The ACF of AR(1):

We can rewrite the formula (3) as:

$$X_t - \mu = \phi_1\left(X_{t-1} - \mu\right) + Z_t$$

If we times X_t in two sides and take expectation, we know Z_t is independent to sequence $\{X_t\}$, then

$$E\left[X_{t-k}\left(X_t - \mu\right)\right] = \phi_1 E\left[X_{t-k}\left(X_{t-1} - \mu\right)\right] + E\left[Z_t X_{t-k}\right] = \phi_1 E\left[X_{t-k}\left(X_{t-1} - \mu\right)\right]$$

We know that

$$\gamma_k = \operatorname{Cov}\left(X_{t-k}, \left(X_t - \mu\right)\right) = E\left[X_{t-k}\left(X_t - \mu\right)\right]$$

so we can get

$$\gamma_k = \phi_1 \gamma_{k-1}, \forall k > 0$$

so

$$\gamma_k = \phi_1 \gamma_{k-1} = \phi_1^2 \gamma_{k-2} = \cdots = \phi_1^k \gamma_0$$

so

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k$$

In general, $\rho_k = \phi_1^k$ and ACF ρ_k decays exponentially as k increases.

Forecast:

[1] necessary and sufficient condition $|\phi| < 1$

[2] if $\operatorname{Var}\left(X_t\right) > 0$, then $1-\phi_1^2 > 0$, so $\phi_1^2 < 1$.

- 1-step ahead forecast at time n, the forecast origin:

$$\hat{X}_n(1) = E\left(X_{n+1}|\mathscr{F}_n\right) = E\left(\phi_0 + \phi_1 X_n + Z_{n+1}|\mathscr{F}_n\right) = \phi_0 + \phi_1 X_n$$

  then 1-step ahead forecast error:

$$e_n(1) = X_{n+1} - \hat{X}_n(1) = Z_{n+1}$$

  Variance of 1-step ahead forecast error:

$$\mathrm{Var}\left[e_n(1)\right] = \mathrm{Var}\left(Z_{n+1}\right) = \sigma^2$$

- 2-step ahead forecast:

$$\hat{X}_n(2) = E\left(X_{n+2}|\mathscr{F}_n\right) = E\left(\phi_0 + \phi_1 X_{n+1} + Z_{n+2}|\mathscr{F}_n\right) = \phi_0 + \phi_1 \hat{X}_n(1)$$

  then 1-step ahead forecast error:

$$\begin{aligned} e_n(2) &= X_{n+2} - \hat{X}_n(2) \\ &= \phi_0 + \phi_1 X_{n+1} + Z_{n+2} - \phi_0 - \phi_1 \hat{X}_n(1) \\ &= \phi_1 X_{n+1} + Z_{n+2} - \phi_1 \hat{X}_n(1) \\ &= \phi_1(\phi_0 + \phi_1 X_n + Z_{n+1}) + Z_{n+2} - \phi_1(\phi_0 + \phi_1 X_n) \\ &= Z_{n+2} + \phi_1 Z_{n+1} \end{aligned}$$

  Variance of 1-step ahead forecast error:

$$\mathrm{Var}\left[e_n(2)\right] = \mathrm{Var}\left(Z_{n+2} + \phi_1 Z_{n+1}\right) = (1 + \phi_1^2)\sigma^2$$

  which is greater than or equal to $\mathrm{Var}\left[e_n(1)\right]$, implying that uncertainty in forecasts increases as the number of step increases.

Compact form:

Formula (3) also can be written as:

$$\left(1 - \phi_1 B\right) X_t = \phi_0 + Z_t$$

where B is lag operator.

2.1.2 AR(2) model

The form of AR(2) can be written as:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t \tag{4}$$

or

$$\left(1 - \phi_1 B - \phi_2 B^2\right) X_t = \phi_0 + Z_t$$

Staionarity condition: all roots of $\left(1 - \phi_1 z - \phi_2 z^2\right) = 0$ lie outside the unit circle.

If we consider reciprocal of root z, we can rewrite the characteristic equation $\left(1 - \phi_1 z - \phi_2 z^2\right) = 0$ as

$$\left(\pi^2 - \phi_1 \pi - \phi_2\right) = 0 \tag{5}$$

where the π is the reciprocal of z. Since $|z| > 1$, we know that $|\pi| < 1$.

Supposed π_1 and π_2 are the roots of (5), then we can get

$$\begin{cases} \pi_1 + \pi_2 &= \phi_1 \\ \pi_1 \pi_2 &= -\phi_2 \end{cases}$$

While

$$\begin{aligned} \phi_1 + \phi_2 - 1 &= \pi_1 + \pi_2 - \pi_1\pi_2 - 1 \\ &= (\pi_1 - 1)(1 - \pi_2) < 0 \end{aligned}$$

and

$$\begin{aligned} \phi_1 - \phi_2 - 1 &= -\pi_1\pi_2 - \pi_1 - \pi_2 - 1 \\ &= -(\pi_1 + 1)(\pi_2 + 1) < 0 \end{aligned}$$

so we get

$$\left\{ \begin{array}{l} \phi_1 + \phi_2 - 1 < 0 \\ \phi_1 - \phi_2 - 1 < 0 \end{array} \right. \longrightarrow \phi_2 < 1 \pm \phi_1$$

Since $|\pi_1| < 1, |\pi_2| < 1$, we can get $|\phi_2| = |\pi_1\pi_2| < 1$, thus

$$\left\{ \begin{array}{c} |\phi_2| < 1 \\ \phi_2 < 1 \pm \phi_1 \end{array} \right.$$

Stochastic business cycle: As the characteristic equation is

$$1 - \phi_1 z - \phi_2 z^2 = 0,$$

the discriminant of root is $\Delta = \phi_1^2 + 4\phi_2$.

If $\phi_1^2 + 4\phi_2 < 0$, the root is complex. The roots can be written as:

$$z_1, z_2 = \frac{\phi_1}{-2\phi_2} \pm i\frac{\sqrt{|\phi_1^2 + 4\phi_2|}}{-2\phi_2}$$

then the modulus is

$$|z_i| = \frac{1}{\sqrt{-\phi_2}}$$

Since $x = |x|(\cos\theta + i\sin\theta)$, then radian is

$$\theta = \cos^{-1}\frac{\phi_1}{2\sqrt{-\phi_2}}$$

so the cycle of this radian is

$$k = \frac{2\pi}{\cos^{-1}\phi_1/(2\sqrt{-\phi_2})}$$

If we denote the solutions of the polynomial as $a \pm bi$, where $i = \sqrt{-1}$, then

$$x = a \pm bi = \sqrt{a^2 + b^2}(\cos\theta + i\sin\theta)$$

so that

$$k = \frac{2\pi}{\cos^{-1}\left(a/\sqrt{a^2 + b^2}\right)}$$

The ACF of AR(2):

The (4) can be rewritten as

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + Z_t$$

If we times X_{t-k} in two sides and take expectation, then

$$\begin{aligned} E\left[X_{t-k}\left(X_t - \mu\right)\right] &= \phi_1 E\left[X_{t-k}\left(X_{t-1} - \mu\right)\right] + \phi_2\left[X_{t-k}(X_{t-2} - \mu)\right] + E\left[Z_t X_{t-k}\right] \\ &= \phi_1 E\left[X_{t-k}\left(X_{t-1} - \mu\right)\right] + \phi_2\left[X_{t-k}(X_{t-2} - \mu)\right] \\ &= \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} \end{aligned}$$

so that

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}, \quad k \geq 2$$

Forecast: Similar to AR(1) model.

2.2 Moving Average (MA) model

2.2.1 MA(1) model

The form of MA(1) model is:

$$X_t = \mu + Z_t + \theta_1 Z_{t-1} \tag{6}$$

Expectation:

$$E(X_t) = \mu$$

Variance:

$$\operatorname{Var}(r_t) = \left(1 + \theta_1^2\right)\sigma^2$$

Covariance:

$$\gamma_1 = Cov(X_t, X_{t-1}) = E\left[(X_t - \mu)(X_{t-1} - \mu)\right] = E\left[(Z_t + \theta_1 Z_{t-1})(Z_{t-1} + \theta_1 Z_{t-2})\right] = \theta_1 E(Z_{t-1}^2) = \sigma^2\theta_1$$

As for $\forall k > 1$,

$$\gamma_k = E\left[(Z_t + \theta_1 Z_{t-1})(Z_{t-k} + \theta_1 Z_{t-k-1})\right] = 0$$

so that

$$\gamma_k = \begin{cases} \sigma^2\left(1 + \theta_1^2\right), & k = 0 \\ \sigma^2\theta_1, & k = 1 \\ 0, & k > 1 \end{cases}$$

so the ACF of MA(1) model is

$$\rho_k = \begin{cases} 1, & k = 0 \\ \frac{\theta_1}{1+\theta_1^2}, & k = 1 \\ 0, & k > 1 \end{cases}$$

Forecast:

- 1-step ahead: (at origin $t = n$)

  $$\hat{X}_n(1) = E\left(X_{n+1} | X_1, \ldots, X_n\right) = \mu + \theta_1 Z_n$$

  1-step ahead forecast error:

  $$e_n(1) = Z_{n+1}$$

  with variance σ^2.

- multi-step ahead:

  $$\hat{X}_n(k) = \mu$$

  for $k > 1$.

Invertibility:

- Concept: X_t is a proper linear combination of Z_t and the past observations $\{X_{t-1}, X_{t-2}, \cdots\}$.
- For an invertible model, the dependence of X_t on X_{t-k} converges to zero as k increases.
- Condition: $|\theta| < 1$
- Invertibility of MA models is the dual property of stationarity for AR models.

2.2.2 MA(2) model

The form of MA(2) can be written as:

$$X_t = \mu + Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

or

$$X_t - \mu = (1 + \theta_1 B + \theta_2 B^2) Z_t$$

Invertibility: all the roots of $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 = 0$ lie outside the unit circle.

2.3 ARMA(p,q)

The ARMA(p,q) model can be represented as:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \tag{7}$$

where $Z_t \sim N(0, \sigma^2)$.

3 Seasonal time series

3.1 Multiplicative model

The form can be represented as:

$$X_t - X_{t-1} - X_{t-4} + X_{t-5} = Z_t - \theta_1 Z_{t-1} - \theta_4 Z_{t-4} + \theta_1\theta_4 Z_{t-5}$$

or

$$(1-B)\left(1-B^4\right) X_t = (1-\theta_1 B)\left(1-\theta_4 B^4\right) Z_t$$

Supposed that

$$Y_t = (1-\theta_1 B)\left(1-\theta_4 B^4\right) Z_t = Z_t - \theta_1 Z_{t-1} - \theta_4 Z_{t-4} + \theta_1\theta_4 Z_{t-4-1}$$

then

- $E(Y_t) = 0$
- $\gamma_0 = \operatorname{Var}(Y_t) = \left(1+\theta_1^2\right)\left(1+\theta_4^2\right)\sigma^2$
- $\gamma_1 = -\theta_1\left(1+\theta_4^2\right)\sigma^2$
- $\gamma_3 = \theta_1\theta_4\sigma^2$
- $\gamma_4 = -\theta_4\left(1+\theta_1^2\right)\sigma^2$
- $\gamma_5 = \theta_1\theta_4\sigma^2$
- $\gamma_k = 0, k \neq 0, 1, 3, 4, 5$

so the ACF of $\{Y_t\}$ is

- $\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$
- $\rho_4 = \frac{-\theta_4}{1+\theta_4^2}$
- $\rho_3 = \rho_5 = \rho_1\rho_4 = \frac{\theta_1\theta_4}{\left(1+\theta_1^2\right)\left(1+\theta_4^2\right)}$
- $\rho_k = 0$, for other k

As for

$$(1-B)\left(1-B^s\right) X_t = (1-\theta B)\left(1-\Theta B^s\right) Z_t$$

we can get covariance as

- $E(Y_t) = 0$
- $\gamma_0 = \operatorname{Var}(Y_t) = \left(1+\theta^2\right)\left(1+\Theta^2\right)\sigma^2$
- $\gamma_1 = -\theta\left(1+\Theta^2\right)\sigma^2$
- $\gamma_{s-1} = \theta\Theta\sigma^2$
- $\gamma_s = -\Theta\left(1+\theta^2\right)\sigma^2$
- $\gamma_{s+1} = \theta\Theta\sigma^2$

- $\gamma_k = 0, k \neq 0, 1, s-1, s, s+1$

so the ACF of $\{Y_t\}$ becomes

- $\rho_1 = \frac{-\theta}{1+\theta^2}$
- $\rho_s = \frac{-\Theta}{1+\Theta^2}$
- $\rho_{s-1} = \rho_{s+1} = \rho_1\rho_s = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)}$
- $\rho_k = 0$, for other k

4 GARCH models - Condtional heteroscedastic models

4.1 ARCH effect

Let $\xi_t = Z_t^2 - EZ_t^2$. If there is not ARCH effect, then the ACF of ξ_t are all zero.
In 1983, McLeod and Li use Ljung-Box to test the null H_0: the ACF ρ_k of ξ_t are all zero, i.e,

$$H_0 : \rho_1 = \cdots = \rho_m = 0$$

Let $\hat{\rho}_k$ be the sample ACF of $\{\xi_t\}$. We use the Ljung-Box:

$$Q(m) = n(n-1)\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k} \sim \chi^2(m)$$

If test statistics is significant, then reject H_0 and there exists ARCH effect; otherwise, we can not reject the null hypothesis, which means there do not exist ARCH effect.

In 1982, Engle use (Lagrange multiplier) LM test to test ARCH effect: Assumed that

$$\xi_t = \alpha_1\xi_{t-1} + \cdots + \alpha_m\xi_{t-m} + e_t = \beta' U_t + e_t$$

where $\beta = (\alpha_1, \cdots, \alpha_m)'$ and $X_t = (\xi_{t-1}, \cdots, \xi_{t-m})$ and use LM test for the null:

$$H_0 : \alpha_1 = \cdots = \alpha_m = 0$$

The LM test statistics:

$$LM = \left(\sum_{t=1}^{n} X_t'\right)\left[\sum_{t=1}^{n} X_t X_t'\right]^{-1}\left(\sum_{t=1}^{n} X_t\right) \sim \chi^2(m)$$

4.2 ARCH model

$$Z_t = \sigma_t\epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2 + \cdots + \alpha_m Z_{t-m}^2$$

where $\{\epsilon_t\}$ is a sequence of iid r.v. with mean 0 and variance 1, also $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$.

4.3 GARCH model

$$Z_t = \sigma_t\epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m}\alpha_i Z_{t-i}^2 + \sum_{i=1}^{s}\beta_i\sigma_{t-i}^2$$

where $\{\epsilon_t\}$ is defined above, also $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$.

Focus on a GARCH(1,1) model:

$$Z_t = \sigma_t\epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

In order to calculate expectation $E(Z_t)$, then firstly calculate conditional expectaion:

$$
\begin{aligned}
E\left(Z_t|F_{t-1}\right) = E\left(\sigma_t\epsilon_t|F_{t-1}\right) &= E\left(\sqrt{\alpha_0+\alpha_1 Z_{t-1}^2+\beta_1\sigma_{t-1}^2}\cdot\epsilon_t|F_{t-1}\right)\\
&= \sqrt{\alpha_0+\alpha_1 Z_{t-1}^2+\beta_1\sigma_{t-1}^2}\cdot E\left(\epsilon_t|F_{t-1}\right)\\
&= \sigma_t E\left(\epsilon_t|F_{t-1}\right) = 0
\end{aligned}
$$

so that

$$E(Z_t) = E\left[E\left(Z_t|F_{t-1}\right)\right] = 0$$

Next, we try to calculate the variance of Z_t. Supposed that sequence $\{Z_t\}$ exists strictly stationary solution, since

$$
\begin{aligned}
E\left(Z_t^2|F_{t-1}\right) = E\left(\sigma_t^2\epsilon_t^2|F_{t-1}\right) &= E\left((\alpha_0+\alpha_1 Z_{t-1}^2+\beta_1\sigma_{t-1}^2)\epsilon_t^2|F_{t-1}\right)\\
&= (\alpha_0+\alpha_1 Z_{t-1}^2+\beta_1\sigma_{t-1}^2)E\left(\epsilon_t^2|F_{t-1}\right)\\
&= \sigma_t^2 E\left(\epsilon_t^2|F_{t-1}\right)\\
&= \sigma_t^2
\end{aligned}
$$

then

$$
\begin{aligned}
\operatorname{Var}\left(Z_t\right) &= E\left(Z_t^2\right)\\
&= E\left[E\left(Z_t^2|F_{t-1}\right)\right]\\
&= E\left[E\left(\sigma_t^2\epsilon_t^2|F_{t-1}\right)\right]\\
&= E\left[\sigma_t^2 E\left(\epsilon_t^2|F_{t-1}\right)\right]\\
&= E\left[\sigma_t^2 E\left(\epsilon_t^2\right)\right]\\
&= E\left[\sigma_t^2\right]\\
&= E\left[\alpha_0+\alpha_1 Z_{t-1}^2+\beta_1\sigma_{t-1}^2\right]\\
&= \alpha_0+\alpha_1 E\left(Z_{t-1}^2\right)+\beta_1 E\left(\sigma_{t-1}^2\right)\\
&= \alpha_0+\alpha_1 E\left(Z_{t-1}^2\right)+\beta_1 E\left[E\left(Z_{t-1}^2|F_{t-2}\right)\right]\\
&= \alpha_0+\left(\alpha_1+\beta_1\right)E\left(Z_{t-1}^2\right)
\end{aligned}
$$

Let $E(Z_t^2)=E(Z_{t-1}^2)$, then we get

$$\operatorname{Var}\left(Z_t\right) = E(Z_t^2) = \frac{\alpha_0}{1-\alpha_1-\beta_1}$$

4.3.1 Forecast

Since

$$\sigma_{t+1}^2 = \alpha_0+\alpha_1 Z_t^2+\beta_1\sigma_t^2 \in F_t$$

then 1-step ahead forescast:

$$\sigma_t^2(1) = E\left(\sigma_{t+1}^2|F_t\right) = \sigma_{t+1}^2 = \alpha_0+\alpha_1 Z_t^2+\beta_1\sigma_t^2$$

As for 2-step ahead forecast, since

$$Z_t^2 = \sigma_t^2\epsilon_t^2$$

and

$$
\begin{aligned}
\sigma_{t+2}^2 &= \alpha_0+\alpha_1 Z_{t+1}^2+\beta_1\sigma_{t+1}^2\\
&= \alpha_0+\alpha_1\sigma_{t+1}^2\epsilon_{t+1}^2+\beta_1\sigma_{t+1}^2\\
&= \alpha_0+\left(\alpha_1\epsilon_{t+1}^2+\beta_1\right)\sigma_{t+1}^2
\end{aligned}
$$

then

$$\sigma_t^2(2) = E\left(\sigma_{t+2}^2|F_t\right) = \alpha_0+E\left(\alpha_1\epsilon_{t+1}^2+\beta_1|F_t\right)\sigma_{t+1}^2 = \alpha_0+\left(\alpha_1+\beta_1\right)\sigma_t^2(1)$$

In general, since

$$\sigma_{t+\ell}^2 = \alpha_0+\alpha_1\varepsilon_{t+\ell-1}^2\sigma_{t+\ell-1}^2+\beta_1\sigma_{t+\ell-1}^2 = \alpha_0+\left(\alpha_1\varepsilon_{t+\ell-1}^2+\beta_1\right)\sigma_{t+\ell-1}^2$$

we have

$$\begin{aligned}\sigma_t^2(\ell) &= E\left\{\sigma_{t+\ell}^2 | F_t\right\} = \alpha_0 + E\left\{\left(\alpha_1\epsilon_{t+\ell-1}^2 + \beta_1\right)\sigma_{t+\ell-1}^2 | F_t\right\} \\ &= \alpha_0 + E\left\{E\left[\left(\alpha_1\epsilon_{t+\ell-1}^2 + \beta_1\right)\sigma_{t+\ell-1}^2 | F_{t+\ell-2}\right] | F_t\right\} \\ &= \alpha_0 + E\left\{\sigma_{t+\ell-1}^2 E\left[\alpha_1\varepsilon_{t+\ell-\ell-1}^2 + \beta_1 | F_{t+\ell-2}\right] | F_t\right\} \\ &= \alpha_0 + \left\{\sigma_{t+\ell-1}^2\left(\alpha_1 + \beta_1\right) | F_t\right\} \\ &= \alpha_0 + \left(\alpha_1 + \beta_1\right)\sigma_t^2(\ell - 1)\end{aligned}$$

4.4 IGARCH model

An IGARCH model is as follows:

$$Z_t = \sigma_t\epsilon_t, \sigma_t^2 = \alpha_0 + \beta_1\sigma_{t-1}^2 + \left(1 - \beta_1\right)Z_{t-1}^2$$

In this case,

$$\sigma_t^2(\ell) = \sigma_t^2(1) + (\ell - 1)\alpha_0$$

where h is the forecast origin. The effect of $\sigma_t^2(\ell)$ on future is persistent, and the volatility forecasts form a straight line with slope α_0.

4.5 GARCH-M model

$$\begin{gathered} r_t = \mu + c\sigma_t^2 + Z_t \\ Z_t = \sigma_t\epsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2 + \beta_1\sigma_{t-1}^2 \end{gathered}$$

where c is referred to as **risk premium**, which is expected to be positive.

4.6 EGARCH model

Asymmetry in responses to + and - returns:

$$g\left(\epsilon_t\right) = \theta\epsilon_t + \left[\left|\epsilon_t\right| - E\left(\left|\epsilon_t\right|\right)\right]$$

with $E\left[g\left(\epsilon_t\right)\right] = 0$. To see asymmetry of $g\left(\epsilon_t\right)$, rewrite it as

$$g\left(\epsilon_t\right) = \begin{cases} (\theta + 1)\epsilon_t - E\left(\left|\epsilon_t\right|\right) & \text{if } \epsilon_t \geq 0 \\ (\theta - 1)\epsilon_t - E\left(\left|\epsilon_t\right|\right) & \text{if } \epsilon_t < 0 \end{cases}$$

An EGARCH(m,s) model is formulated as:

$$Z_t = \sigma_t\epsilon_t, \ln\left(\sigma_t^2\right) = \alpha_0 + \sum_{i=1}^{m}\alpha_i g\left(\epsilon_{t-i}\right) + \sum_{i=1}^{s}\beta_i\ln\left(\sigma_{t-i}^2\right)$$

Consider an EGARCH(1,1) model

$$Z_t = \sigma_t\epsilon_t, \quad (1 - \beta B)\ln\left(\sigma_t^2\right) = \alpha_0 + \alpha g\left(\epsilon_{t-1}\right)$$

Under normality, $E\left(\left|\epsilon_t\right|\right) = \sqrt{2/\pi}$ and the model becomes

$$(1 - \beta B)\ln\left(\sigma_t^2\right) = \begin{cases} \alpha_* + \alpha(\theta + 1)\epsilon_{t-1} & \text{if } \epsilon_{t-1} \geq 0 \\ \alpha_* + \alpha(\theta - 1)\epsilon_{t-1} & \text{if } \epsilon_{t-1} < 0 \end{cases}$$

where $\alpha_* = \alpha_0 - \alpha\sqrt{2/\pi}$.

4.7 TGARCH model

The Threshold GARCH (TGARCH) or GJR Model A TGARCH(s,m) or GJR(s,m) model is defined as

$$\begin{gathered} r_t = \mu_t + Z_t, \quad Z_t = \sigma_t\epsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{s}\left(\alpha_i + \gamma_i N_{t-i}\right)Z_{t-i}^2 + \sum_{j=1}^{m}\beta_j\sigma_{t-j}^2 \end{gathered}$$

where N_{t-i} is an indicator variable such that

$$N_{t-i} = 1 \text{ if } Z_{t-i} < 0, \text{ and } = 0 \text{ otherwise}$$

5 Multivariate time series models

The real data Z_t is a vector:

$$Z_t = \begin{bmatrix} Z_{1t} \\ Z_{2t} \\ \vdots \\ Z_{mt} \end{bmatrix}$$

Z_t is called an m-dimensional vector time series.

- Mean
$$EZ_t = \begin{bmatrix} EZ_{1t} \\ EZ_{2t} \\ \vdots \\ EZ_{mt} \end{bmatrix} = \begin{bmatrix} \mu_{1t} \\ \mu_{2t} \\ \vdots \\ \mu_{mt} \end{bmatrix} \equiv \mu$$

- Covariance
$$\begin{aligned} \Gamma(k) &= \operatorname{Cov}\left(Z_t, Z_{t+k}\right) = E\left[\left(Z_t - \mu\right)\left(Z_{t+k} - \mu\right)'\right] \\ &= \begin{bmatrix} \gamma_{11}(k) & \gamma_{12}(k) & \cdots & \gamma_{1m}(k) \\ \gamma_{21}(k) & \gamma_{22}(k) & \cdots & \gamma_{2m}(k) \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{m1}(k) & \gamma_{m2}(k) & \cdots & \gamma_{mm}(k) \end{bmatrix} \end{aligned}$$

- Corralation matrix
$$\rho_0 = \left(\rho_{ij}(0)\right)_{k\times k} = D^{-1}\Gamma_0 D^{-1}$$
where $D = \operatorname{diag}\left(\sqrt{\Gamma_{11}(0)}, \ldots, \sqrt{\Gamma_{kk}(0)}\right)$

As for k-dimension time series,

$$\begin{aligned} \Gamma_{ij}(l) &= \operatorname{Cov}\left(r_{it}, r_{j,t-l}\right) = \operatorname{Cov}\left(r_{j,t-l}, r_{i,t}\right) = \operatorname{Cov}\left(r_{j,t}, r_{i,t+l}\right) \\ &= \operatorname{Cov}\left(r_{j,t}, r_{i,t-(-l)}\right) = \Gamma_{ji}(-l) \end{aligned}$$

so that

$$\Gamma_{-l} = \Gamma_l^T$$

and also we can get

$$\rho_{-l} = \rho_l^T$$

5.1 The vector AR(1) model

- Model
$$\left(I - \Phi_1 B\right)\dot{X}_t = Z_t, \quad \text{or} \dot{X}_t = \Phi_1 \dot{X}_{t-1} + Z_t$$
when $m = 2$,
$$\begin{bmatrix} \dot{X}_{1t} \\ \dot{X}_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} \dot{X}_{1,t-1} \\ \dot{X}_{2,t-1} \end{bmatrix} + \begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix}$$

- Stationarity conditions

  All the roots of the determinant $|I - \Phi_1 z| = 0$ lie outside the unit cricle.

  In this case,
$$\left(I - \Phi_1 B\right)^{-1} = I + \Phi_1 B + \Phi_1^2 B^2 + \cdots$$
  and
$$\dot{X}_t = Z_t + \Phi_1 Z_{t-1} + \Phi_1^2 Z_{t-2} + \cdots$$

- Covariance matrix function

$$\Gamma(k) = \begin{cases} \Gamma(-1)\Phi_1' + \Sigma, & \text{if } k = 0 \\ \Gamma(k-1)\Phi_1' = \Gamma(0)\left(\Phi_1'\right)^k, & \text{if } k \geq 1 \end{cases}$$

In general, $\Gamma(0) \neq I$

$$\Phi_1 = \Gamma'(1)\Gamma^{-1}(0) \quad \text{and } \Sigma = \Gamma(0) - \Gamma'(1)\Phi_1'$$

Thus,

$$\Sigma = \Gamma(0) - \Phi_1\Gamma(0)\Phi_1'$$

We can find $\Gamma(0)$ by solving the above equation.

5.2 The vector MA(1)model

- Model

$$\dot{Z}_t = (I - \Theta_1 B)\, a_t, \quad \text{or } \dot{Z}_t = a_t - \Theta_1 a_{t-1}$$

when $m = 2$,

$$\begin{bmatrix} \dot{Z}_{1t} \\ \dot{Z}_{2t} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}$$

- Invertibility conditions

All the roots of the determinant $|I - \Theta_1 z| = 0$ lie outside the unit circle.

In this case,

$$(I - \Theta_1 B)^{-1} = I + \Theta_1 B + \Theta_1^2 B^2 + \cdots$$

and

$$Z_t = \dot{X} + \Theta_1 \dot{X}_{t-1} + \Theta_1^2 \dot{X}_{t-2} + \cdots$$

- Covariance matrix function

$$\Gamma(0) = \Sigma + \Theta_1 \Sigma \Theta_1';$$

$$\Gamma(k) = \begin{cases} -\Sigma\Theta_1', & \text{if } k = 1 \\ -\Theta_1\Sigma, & \text{if } k = -1 \\ 0, & \text{if } |k| > 1 \end{cases}$$

6 Partially non-stationary vector time series

6.1 Cointegration

Definition 1. (Engle and Granger 1987): For a series yt with no deterministic components, let

$$(1 - B)^d y_t = z_t$$

If z_t is a stationary ARMA process, we said that $y_t \sim I(d)$.

Definition 1. (Engle and Granger 1987): If all elements of the vector $\mathbf{y}_t$ are $I(d)$ and there exists a vector $\beta \neq 0$ such that

$$\beta' \mathbf{y}_t \sim I(d - b)$$

for some $b > 0$, the vector process is said to be cointegrated $CI(d, b)$ and β is called cointegrating vector.

Example. The bivariate system:

$$y_{1t} = \gamma y_{2t} + \varepsilon_{1t}$$
$$y_{2t} = y_{2\{t-1\}} + \varepsilon_{2t}$$

where $\gamma \neq 0$, ε_{1t} and ε_{2t} being uncorrelated white noise processes.

$\triangle y_{2t} = \varepsilon_{2t}$, where $\triangle \equiv 1 - B$.

$$\triangle y_{1t} = \gamma \triangle y_{2t} + \triangle \varepsilon_{1t} = \gamma \varepsilon_{2t} + \varepsilon_{1t} - \varepsilon_{1,t-1}$$

Thus, both y_{1t} and y_{2t} are $I(1)$ processes, but the linear combination $y_{1t} - \gamma y_{2t}$ is stationary.

Hence $\mathbf{y}_t = (y_{1t}, y_{2t})'$ is cointegrated with a cointegrating vector $\beta = (1, -\gamma)'$.

In general, if the vector process $\mathbf{y}_t$ has m components, then there can be more than one cointegrating vector β'.