第十讲：非平稳序列：异方差模型（二）

概要： 本讲介绍 GARCH 模型的建模步骤。主要内容包括 PP 检验、异方差检验、拟合后检验和预测。

我们拿到一个观测值序列，**完整的分析应该同时关注序列均值和序列波动两方面信息的提取**。通常先提取序列的均值信息，然后分析残差序列中蕴涵的波动信息。将这两方面信息结合起来，才是比较完整和精确的分析结果。一个完整的模型如下：

$$\begin{cases} X_t = f(t, X_{t-1}, X_{t-2}, \ldots) + \varepsilon_t \\ \varepsilon_t = e_t\sqrt{h_t} \\ h_t = \lambda_0 + \sum_{j=1}^{p}\eta_j h_{t-j} + \sum_{i=1}^{q}\lambda_i \varepsilon_{t-i}^2 \end{cases} \qquad 9.1$$

模型包含三个式子，第一个式子定义了$\{x_t\}$确定性信息的拟合模型，或者说是均值的拟合模型，实现残差序列零均值，且相关信息提取充分，即$E(\varepsilon_t) = 0$,且$\rho_k = 0$。在本学期的讲授中，通常是 ARMA/ARIMA 模型，第二个式子中的e_t为原序列提取确定性信息和条件异方差信息之后的残差波动。因为序列均值和方差中的相关信息都提取干净了，所以e_t应该是真正的白噪声序列，同时依据e_t的特征定义了残差项及序列的分布类型，通常假定e_t服从正态分布。第三个式子定义了条件异方差模型，通过构造残差平方的 ARMA 模型，得到了 GARCH(p,q)模型，其中，当$p = 0$时，为 ARCH(q)模型。

拟合 GARCH 模型的步骤如下：

第一步：构建水平模型$x_t = f(t, x_{t-1}, x_{t-2}, \ldots) + \varepsilon_t$，提取序列均值中蕴涵的相关信息；

第二步：检验残差序列$\{\varepsilon_t\}$是否具有条件异方差特征

第三步： 对具有条件异方差特征的序列拟合 GARCH 模型

第四步：检验拟合模型的优劣，优化模型

第五步：使用拟合模型进行预测

在 ARMAARIMA 模型的建模步骤中，ADF 检验统计量和白噪声检验 LB / Q 统计量的适用前提是同方差，拟合后的模型检验，以残差项序列｛ε_t｝是否为白噪声作为步骤的结束，也是以同方差为假设前提的，但是否同方差，必须进行检验，同时，如果存在异方差，平稳性检验和白噪声检验的方法要改变。

二、PP 检验

建立条件异方差模型，首先需要提取序列均值（确定性）信息，最常用的均值模型是 ARMA/ARMA 模型。要建立 ARMA/ARIMA 模型，必须适用 ADF 检验，对序列的平稳性进行判断。使用 ADF 检验有一个基本假定：

$$Var(\varepsilon_t)=\sigma^2$$

即 ADF 检验只适用于方差齐性场合，它对异方差序列的平稳性检验就可能出现偏差。Phillips 和 Perron 于 1988 年对 ADF 检验进行了非参数修正，提出了 PP 检验，即异方差条件下的平稳性检验，这里的平稳性检验指的是检验均值平稳。

PP 检验，残差序列$\{\varepsilon_t\}$需要满足以下三个条件：

(1) 均值为零，即$E(\varepsilon_t)=0$

(2) 方差有至少一个高阶矩存在，即$supE(|\varepsilon_t|^2)<0$，且对于某个$\beta>2$，有$supE(|\varepsilon_t|^\beta)<0$，由于没有假定$E(|\varepsilon_t|^2)$为常数，这个条件实际上意味着允许异方差存在。

(3) 非退化极限分布存在，即$\sigma_s^2=\lim\limits_{T\to\infty}[E({S_T}^2/T)]$存在且为正值，其中，T 为序列长度，$\sum_{t=1}^{T}\varepsilon_t=S_T$。

以下以 AR(1)模型$X_t=\phi_1X_{t-1}+\varepsilon_t$ 为例，介绍 PP 检验的构造原理。

假定$\widehat{\phi_1}$是ϕ_1的最小二乘估计值，那么$\widehat{\phi_1}$的方差可以定义为：

$$\sigma_{\widehat{\phi_1}}^2=\lim_{T\to\infty}\left[\sum_{t=1}^{T}E(\varepsilon_t^2)\Big/T\right] \qquad 9.2$$

当$\{\varepsilon_t\}$为白噪声序列时，有

$$\sigma_{\widehat{\phi_1}}^2=\sigma_s^2 \qquad 9.3$$

如果$\{\varepsilon_t\}$不是白噪声序列，则式 9.3 不成立。为了说明式 9.3 不成立，假定$\{\varepsilon_t\}$服从 MA(1)模型，即；

$$\varepsilon_t=\alpha_t-\theta_1\alpha_{t-1}$$

$$\sigma_{\widehat{\phi_1}}^2=E(\varepsilon_t^2)=var(\varepsilon_t)=(1+\theta_1^2)\sigma_\alpha^2$$

而

$$\sigma_s^2=E(\varepsilon_1^2)+2\sum_{j=2}^{\infty}E(\varepsilon_1\varepsilon_j)=(1+\theta_1^2-2\theta_1)\sigma_\alpha^2$$

显然，$\sigma_{\phi_1}^2 \neq \sigma_S^2$。Phillips 和 Perron 正是利用这种不等，使用$\sigma_{\phi_1}^2$和σ_S^2的估计值对 ADF 检验的τ统计量进行了非参数修正，修正后的统计量如下：

$$Z(\tau) = \tau\left(\widehat{\sigma_{\phi_1}}^2 / \widehat{\sigma_{Sl}}^2\right) - \frac{1}{2}\left(\widehat{\sigma_{Sl}}^2 - \widehat{\sigma_{\phi_1}}^2\right)T\sqrt{\widehat{\sigma_{Sl}}^2 \sum_{t=2}^{T}\left(x_{t-1} - \overline{x_{T-1}}\right)^2} \qquad 9.4$$

式中：(1) $\widehat{\sigma_{\phi_1}}^2$是$\sigma_{\phi_1}^2$的无条件方差的样本估计值，即：

$$\widehat{\sigma_{\phi_1}}^2 = \sum_{t=1}^{T}\hat{\varepsilon}_t^2 \Big/ T \qquad 9.5$$

(2) 假设可以估计$\{\varepsilon_t\}$显著自相关的延迟结束为 l，$\widehat{\sigma_{Sl}}^2$是σ_S^2的条件方差的样本估计值：

$$\widehat{\sigma_{Sl}}^2 = \sum_{t=1}^{l}\hat{\varepsilon}_t^2 \Big/ T + 2\sum_{j=1}^{t}\phi_j(l)\sum_{t=j+1}^{T}\hat{\varepsilon}_t\widehat{\varepsilon_{t-j}} \qquad 9.6$$

式中，$\phi_j(l) = 1 - [1/(l+1)]$，这个权重保证了$\widehat{\sigma_{Sl}}^2$为正值。

(3) $\overline{x_{T-1}} = \sum_{t=1}^{T-1} x_t / (T-1)$

修正后的$Z(\tau)$统计量和τ统计量具有相同的极限分布。这意味着对于异方差序列只需要在原来τ统计量的基础上进行一定的修正，构造出$Z(\tau)$统计量。$Z(\tau)$统计量不仅考虑到自相关误差所产生的影响，还可以继续使用τ统计量的临界值进行检验，而不需要拟合新的临界指表。

三、条件异方差检验

提取了均值信息后，需要对随机序列进行条件异方差检验。常用的两种检验是 Portmanteau Q 检验和 LM 检验（Lagrange Multiplier Test，拉格朗日乘子检验）。

3.1 Portmanteau Q 检验

这个方法就是之前的检验白噪声的统计量。该检验的构造思路是：如果残差序列具有异方差性，那么**残差平方序列具有自相关性**，即**残差平方序列为非白噪声序列**。也就是说，异方差性检验可以转为残差平方序列的自相关检验，即为残差平方序列的 LB 统计量。

Portmanteau Q 检验的假设条件为：

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_k = 0 \qquad H_1: \rho_1, \rho_2, \ldots, \rho_k \text{不全为零}$$

ρ_k为残差平方序列$\{\varepsilon_t^2\}$的延迟 k 阶自相关系数。统计量的表达式为：

$$Q(q)=n(n+2)\sum_{i=1}^{q}\frac{{\rho_i}^2}{n-i} \qquad 9.7$$

式中，n 为观测值序列长度，ρ_i为ε_t^2序列延迟i阶的自相关系数，有

$$\rho_i=\sqrt{\frac{\sum_{t=i+1}^{n}(\varepsilon_t^2-\overline{\varepsilon^2})(\varepsilon_{t-i}^2-\overline{\varepsilon^2})}{\sum_{t=1}^{n}(\varepsilon_t^2-\overline{\varepsilon^2})^2}} \qquad \overline{\varepsilon^2}=\frac{\sum_{t=1}^{n}\varepsilon_t^2}{n}$$

原假设成立时，Portmanteau Q 检验统计量近似服从自由度为$q-1$的χ^2分布。当检验统计量的 P 值小于显著性水平α时，拒绝原假设，认为残差平方序列具有自相关关系，残差序列存在异方差。

3.2 LM 检验（Lagrange Multiplier Test）

在古典线性回归模型中，其中一个假设就是同方差假设，对于异方差的检验统计量，其中之一就是 LM 检验。在这里，也可以对$\{\varepsilon_t^2\}$进行 LM 检验，以判定是否存在异方差。

LM 检验的构造思想是：如果残差序列方差非齐，且具有集群效应，那么残差平方序列具有自相关性。可以尝试使用自回归模型拟合残差平方序列

$$\varepsilon_t^2=\omega+\sum_{j=1}^{q}\lambda_j\,\varepsilon_{t-j}^2+\alpha_t \qquad 9.8$$

方差齐性的假设就可以转化为这个方程是否显著成立的检验。如果式 9.8 显著成立，即至少存在一个参数λ_j非零，就意味着残差平方序列具有自相关性；如果式 9.8 不能显著成立，即$\lambda_j=0 \quad j=1,2,\ldots,q$，就意味着残差平方序列不存在显著相关性，不能拒绝方差齐性假定。所以，LM 检验的实质就是残差平方序列$\{\varepsilon_t^2\}$自回归方程的显著性检验。

LM 检验的假设条件为：

$$H_0:\lambda_1=\lambda_2=\cdots=\lambda_q=0 \qquad H_1:\lambda_1,\lambda_2,\ldots,\lambda_q\text{不全为零}$$

LM 统计量的表达式为：

$$LM(q)=\frac{(SST-SSE)/q}{SSE/(T-2q-1)} \qquad 9.9$$

其中，SST 为总误差平方和，$SST=\sum_{t=q+1}^{T}\varepsilon_t^2$，自由度为$T-q-1$。SSE 为回归方程残差平方和，$SSE=\sum_{t=q+1}^{T}\alpha_t^2$，自由度为$T-2q-1$。SSR 为回归平方和，$SSR=SST-SSE$，自由度为$q$。在截面数据分析中，$LM(q)$检验统计量服从 F 分布，但在时间序列数据分析中，由于自变量具相关性，所以$LM(q)$检验统计量服从自由度为$q-1$的χ^2分布。当检验统计量的 P 值小于显著性水平α时，拒绝原假设，认为残差平方序列具有自相关关系，残差序列

存在异方差。

四、参数估计

GARCH 模型的参数估计通常使用条件最小二乘估计法和极大似然估计法。

4.1 条件最小二乘估计

条件最小二乘估计就是使拟合模型的误差平方和达到最小的参数的估计方法。GARCH(*p*,*q*)模型的误差平方和为：

$$Q=\sum_{t=p+q-1}^{T-1}(\varepsilon_t^2-h_t)^2=\sum_{t=p+q+1}^{T-1}(\varepsilon_t^2-\lambda_0-\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2-\sum_{i=1}^{p}\eta_i\,h_{t-i})^2 \qquad 9.10$$

要使得式 9.10 的值达到最小，就必须$\frac{\partial Q}{\partial\lambda_i}=0,\ i=0,1,2,3\ldots q$，$\frac{\partial Q}{\partial\eta_i}=0,\ i=1,2,3\ldots p$，这样有$p+q+1$个方程，可以联立方程求出$\lambda_0,\lambda_1,\ldots,\lambda_q,\eta_1,\eta_2,\ldots\eta_p$。

4.2 极大似然估计

在确定序列分布假定之后，根据样本信息写出 GARCH(*p*,*q*)模型的似然函数，然后用数值计算方法，求出使得似然函数达到最大的参数值，这就是极大似然估计法 Maximum Likelihood Estimation, MLE 。

在正态分布假定下，GARCH(*p*,*q*)模型的条件分布是：

$$f\left(\varepsilon_t\middle|\varepsilon_{t-1},\varepsilon_{t-2},\ldots,\varepsilon_{t-q}\right)=\frac{1}{\sqrt{2\pi\sigma_t^2}}\exp\left(\frac{-\varepsilon_t^2}{2\sigma_t^2}\right)$$

$$=\frac{1}{\sqrt{2\pi(\lambda_0+\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i\,h_{t-i})}}\exp\left(\frac{-\varepsilon_t^2}{2\left(\lambda_0+\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i\,h_{t-i}\right)}\right) \qquad 9.11$$

似然函数为：

$$L\left(\lambda_0,\lambda_1,\ldots,\lambda_q,\eta_1,\eta_2,\ldots\eta_p\right)=\prod_{t=1}^{n}f\left(\varepsilon_t\middle|\varepsilon_{t-1},\varepsilon_{t-2},\ldots,\varepsilon_{t-q}\right)=\prod_{t=1}^{n}\frac{1}{\sqrt{2\pi\sigma_t^2}}\exp\left(\frac{-\varepsilon_t^2}{2\sigma_t^2}\right)$$

$$=\prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\lambda_0+\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i\,h_{t-i})}}\exp\left(\frac{-\varepsilon_t^2}{2\left(\lambda_0+\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i\,h_{t-i}\right)}\right) \qquad 9.12$$

对数似然函数为：

$$lnL(\lambda_0,\lambda_1,\ldots,\lambda_q,\eta_1,\eta_2,\ldots\eta_p)$$
$$= -\frac{n}{2}\ln(2\pi)$$
$$-\frac{1}{2}\sum_{t=1}^{n}[\sum_{t=1}^{n}\left(\lambda_0+\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i\,h_{t-i}\right)$$
$$+\frac{\varepsilon_t^2}{2(\lambda_0+\sum_{i=1}^{q}\lambda_i\,\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i\,h_{t-i})}] \qquad 9.13$$

这个对数似然函数为超越方程，所以参数$\lambda_0,\lambda_1,\ldots,\lambda_q,\eta_1,\eta_2,\ldots\eta_p$得不到显式解。但是可以通过数值计算的方法，求出使对数似然函数达到最大的未知参数的值，这个值GARCH(p,q)模型的极大似然估计法。

五、拟合检验

GARCH 模型拟合出来以后，需要对它进行拟合检验。检验内容包含以下三个方面：

5.1 参数显著性检验

GARCH 模型拟合出来以后，首先检验该模型中每个参数是否都显著非零。参数显著性检验和 ARIMA 模型的参数显著性检验一样，构造 t 分布检验统计量，在显著性水平取α（一般取 5%，少数时候取 1%或 10%）时，t统计量的 P 值小于α，认为改参数显著非零，保留该参数；反之，参数不显著，可以删除该参数。

5.2 模型显著性检验

条件异方差模型拟合得好不好，取决于它是否将残差序列中蕴涵得异方差信息充分提取出来。利用拟合模型估计出来得异方差$\widehat{h_t}$，对残差序列和残差平方序列分别进行标准化变换，有：

$$e_t=\frac{\varepsilon_t}{\sqrt{\widehat{h_t}}} \qquad e_t^2=\frac{\varepsilon_t^2}{\widehat{h_t}}$$

如果 GARCH 模型拟合得好，(1)，均值模型应该将序列水平相关信息充分提取出来，残差序列消除异方差影响后应该是白噪声序列，即e_t应该是白噪声序列；(2)，方差模型应该将序列的波动信息充分提取出来，残差平方序列在消除异方差影响之后，应该是白噪声序列，即e_t^2也应该是白噪声序列。

有一个特殊情况需要注意:实践经验表明，很多金融时间序列具有尖峰特征（峰态系数特别大），这时拟合的 GARCH 模型通常能提取该序列典型的条件异方差信息，却不一定能

通过白噪声检验。一个重要的原因是，GARCH 模型的参数估计通常是在正态分布假定下进行的，而序列的尖峰特征表明它们不服从正态分布。由于分布假定不对，导致信息提取不够充分。还有其他一些原因也会影响异方差信息的提取，比如，GARCH 模型的线性结构与真实波动特征不符，导致信息提取不充分;或者 GARCH 模型的信息滞后性偏大，导致对波动信息的提取不够敏感，在大幅波动变小幅波动，或者小幅波动变大幅波动的时候，信息延迟可能会出现很大偏差。

5.3 分布检验

在构造 GARCH 模型时，通常默认该序列服从正态分布。模型的检验之一就是要检验这个分布假设对不对。分布的检验方法有两种：图检法和统计量法。

（1）图检法主要有 QQ 图和直方图。

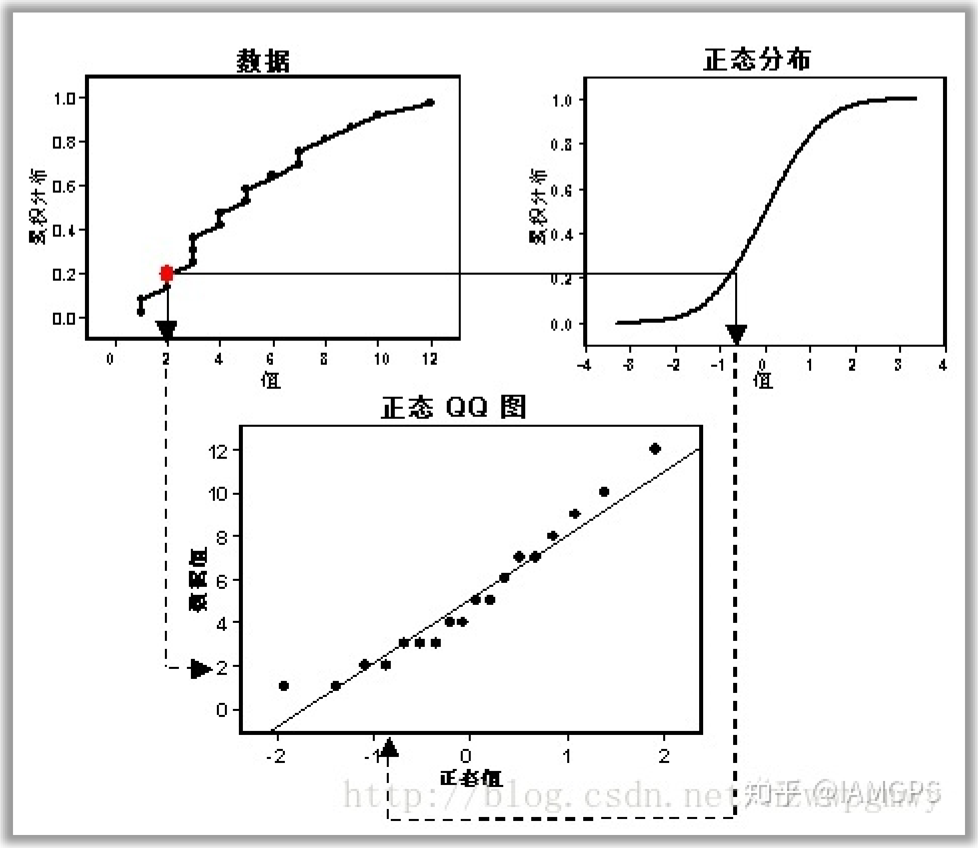

图 10.1 QQ 图的形成

QQ 图全称是 Quantile Quantile Plot，即分位数图。QQ 图是一种散点图，对应于正态分布的 QQ 图，就是由标准正态分布的分位数为横坐标，样本值为纵坐标的散点图。要利用 QQ 图鉴别样本数据是否近似于正态分布，只需看 QQ 图上的点是否近似地在一条直线附近，图形是直线说明是正态分布，用 QQ 图还可获得样本偏度和峰度的粗略信息。

直方图是将残差序列绘制成直方图，然后根据残差序列的样本均值和样本方差，绘制正态分布参考线。如果直方图和正态分布参考线很吻合，说明正态分布的假设比较合理，如果直方图和正态分布参考线的偏离比较严重，说明正态分布的假设不合理。

（2）统计量法。正态性检验常用的统计量有 Jarque-Bera Statistic（JB 检验），Shapiro-Wilk Statistic（SW 检验），Anderson-Darling Statistic（AD 检验）。使用最广泛的是 JB 检验，该统计量构造的思路是：借助正态分布的偏度系数和峰度系数构造出一个服从自由度为 2 的χ^2分布统计量，即：

$$H_0\colon \frac{\varepsilon_t}{\sqrt{h_t}} \sim N(0,1) \qquad H_1\colon \frac{\varepsilon_t}{\sqrt{h_t}}\ not \sim N(0,1)$$

$$JB = \frac{T}{6}b_1^2 + \frac{T}{24}(b_2^2 - 3)^2 \qquad 9.14$$

其中，T为观测值序列的长度，b_1^2为样本的偏度系数，b_2^2为样本的峰度系数。

其实，很多金融时间序列具有尖峰厚尾特征，峰态系数特别大，所以 JB 检验的结果通常会拒绝正态分布假定。这时要不要修正分布假定取决于以下两个条件：

其一，对现有的模型拟合精度满不满意。如果现有模型的拟合精度已经达到研究人员的要求，就可以容忍分布不严格服从正态假定。如果对现有精度不满意，就需要进一步修正分布假定。

其二，有没有适当的分布可替换正态分布。如果分布检验拒绝了正态假定，那么可以用什么新的分布来替代正态分布?新分布是不是比正态分布更好？这需要进一步尝试和研究。通常可以作为正态分布替换分布的是 t 分布、广义误差分布(GED)、椭球等高(elliptical)分布、双曲线(hyperbolic)分布、非对称拉普拉斯(asymmetric Laplace)分布，以及其他具有尖峰厚尾特征的分布。

实践中可能会对一个异方差序列拟合多个 GARCH 模型，每个拟合模型都需要进行上述三方面的拟合检验。除此之外，还需要进行不同模型之间的比较。和 ARIMA 模型的优化一样，我们借助最小信息量准则，比较多个 GARCH 模型的 SBC 或 AIC 的值，信息量最小的拟合模型相对最优。

六、模型预测

假定均值模型为 ARIMA(*p*,*d*,*q*)模型，根据最小均方误差预测原理，使用递推方法得到$t+k$时刻序列均值预测值为$\widehat{x_{t+k}}$，预测方法为：

$$Var(\widehat{x_{t+k}}) = Var(\varepsilon_{t+k}) + \Psi_1^2 Var(\varepsilon_{t+k-1}) + \cdots + \Psi_k^2 Var(\varepsilon_{t-1}) \qquad 9.15$$

在同方差场合：

$$Var(\widehat{x_{t+k}}) = (1 + \Psi_1^2 + \Psi_2^2 \ldots + \Psi_{k-1}^2) Var(\varepsilon_t) \qquad 9.16$$

式 9.16 中，$Var(\varepsilon_t)$为无条件方差，用残差的样本方差估计：

$$\widehat{\sigma_\varepsilon^2} = \frac{\sum_{t=1}^{T} \widehat{\varepsilon_t^2}}{T-1} \qquad 9.17$$

其中，T 为观察值序列的长度，$\widehat{\varepsilon_t^2}$为残差平方的估计值。

在条件异方差场合：

$$Var(\widehat{x_{t+k}}) = \widehat{h_{t+k}} + \Psi_1^2 \widehat{h_{t+k-1}} + \cdots + \Psi_k^2 \widehat{h_{t+1}} \qquad 9.18$$

式中，$\widehat{h_{t+k}}$是根据 GARCH 模型得到的条件异方差预测值。

基于 GARCH 模型，可以使用递推方法得到未来任意期的条件异方差预测值$\widehat{h_{t+k}}$。以 GARCH(1,1)为例，当期（t期）的条件异方差预测值是：

$$\widehat{h_t} = \widehat{\lambda_0} + \widehat{\eta_1} h_{t-1} + \widehat{\lambda_1} \varepsilon_{t-1}^2$$

$t+1$时刻条件异方差的预测值为：

$$\widehat{h_{t+1}} = \widehat{\lambda_0} + \widehat{\eta_1}\widehat{h_t} + \widehat{\lambda_1}\widehat{h_t} = \widehat{\lambda_0} + (\widehat{\eta_1} + \widehat{\lambda_1})\widehat{h_t}$$

以此类推，$t+k$时刻条件异方差的预测值为:

$$\widehat{h_{t+k}} = \widehat{\lambda_0} + \left(\widehat{\eta_1} + \widehat{\lambda_1}\right)\widehat{h_{t+k-1}} \qquad k \geq 1$$

在正态分布假定下，无论是同方差还是条件异方差，该序列的 95%置信区间都是：

$(\widehat{x_{t+k}} - 2\sqrt{Var(\widehat{x_{t+k}})}，\widehat{x_{t+k}} + 2\sqrt{Var(\widehat{x_{t+k}})})$。

七、案例（建模步骤）

对美国经济增长率进行模型拟合：美国季度 GDP（1947 年 1 季度~2012 年 1 季度）

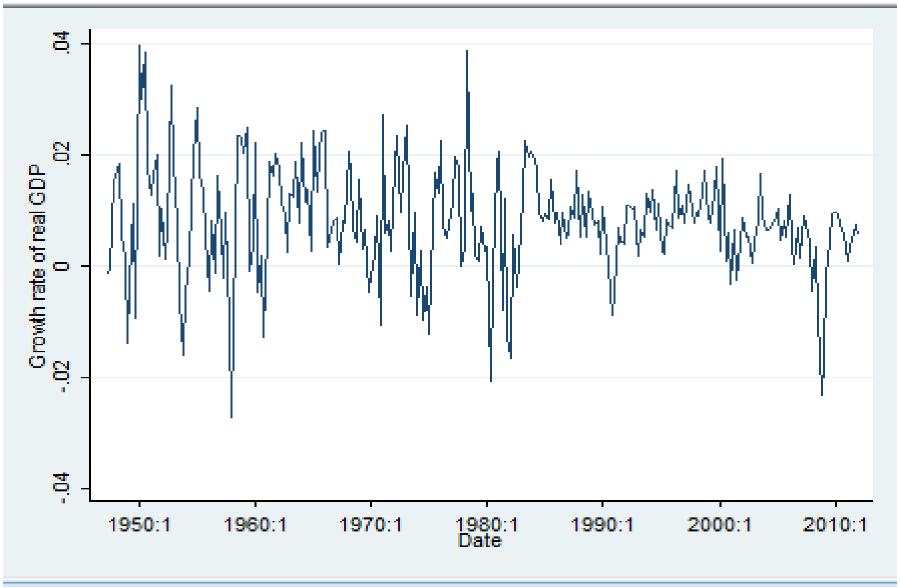

图 10.2 美国季度 GDP 增长率时序图（1947 年 1 季度~2012 年 1 季度）

Step 1：**画出时序图**。从图 10.2 中可以看出，该序列初步判定为均值平稳，但存在一定的集群效应，可能存在异方差；

Step 2：**平稳性检验**。对序列 growth 进行 Unit Root 检验。PP 检验，采用的是类型一，即不包含常数项的回归结构，延迟阶数为 0。计算得到$Z(\tau)$统计量值为-7.836，远小于对应的 1%的临界值。说明原假设 H_0：存在单位根 （即序列非平稳）这个事件发生的概率远小于 0.01。结论：序列 growth 是平稳的。

Step 3：**白噪声检验**。传统的纯随机性检验都采用 LB 统计量，而 LB 统计量是在满足方差齐性的假定下构造的。当序列存在异方差时，LB 统计量不再服从χ^2分布，这说明在异方差条件下，均值的白噪声检验不再准确，此时除了 LB 统计量外，还要参考序列的自相关系数的大小，如果自相关系数绝大部分都小于 0.2，可以认为是白噪声。计算得到 growth 序列的：Portmanteau (Q) statistic = 43.8369，Prob > chi2(6) =0.0000，同时观察序列 growth 的 ACF 图（图 10.3），可以知道 Growth 是非白噪声序列。

Step 4：**模型识别**。计算序列 growth 的自相关系数 ACF（图 10.3）和偏自相关系数 PACF（图 10.4）。从图上看，ACF 拖尾，PACF 截尾，可以用 AR(1)模型进行拟合。

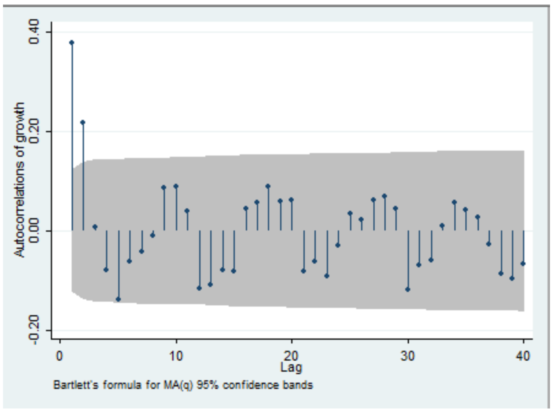

图 10.3 ACF 图

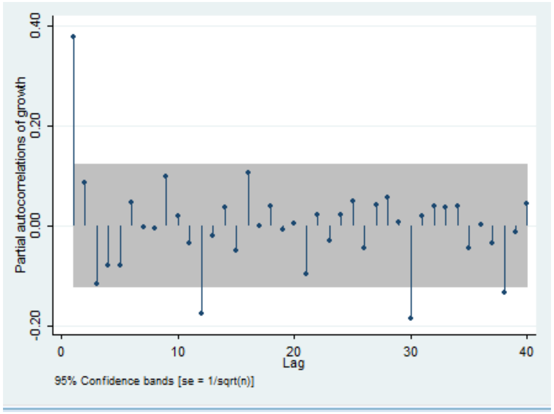

图 10.4 PACF 图

Step 5：**模型估计和参数显著性检验**。任何统计软件在给出模型估计结果时，同时都给出了参数显著性检验。以下以 Stata 13 为例：

```
ARIMA regression

Sample:  1947:2 - 2012:1                        Number of obs      =       260
                                                Wald chi2(1)       =     58.51
Log likelihood =  850.5428                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
      growth |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
growth       |
       _cons |   .0077877   .0009309     8.37   0.000     .0059632    .0096121
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .3762591   .0491894     7.65   0.000     .2798497    .4726685
-------------+----------------------------------------------------------------
      /sigma |   .0091822   .0002939    31.24   0.000     .0086062    .0097583
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
      confidence interval is truncated at zero.
```

从中可以看到，所有参数的 p-value 值都小于 0.05 ，即通过了参数显著性检验。但注意：这个结果是基于同方差假设得到的，因此不是最终的拟合模型。

Step 6：**异方差检验（LM 统计量）**。计算上述 AR(1)模型的残差值，得到残差序列，同时计算残差平方值，得到残差平方序列。利用 LM 统计量，对残差平方项序列进行检验，发现残差平方序列存在自相关性，即存在条件异方差。注意：在 Stata 软件中，进行 LM 检验的同时，也确定了 ARCH 模型的结构（具体可参考 Stata 软件操作教材）。

Step 7：**模型估计和参数显著性检验**。对均值模型 AR(1)和条件异方差模型进行联合模型估计，软件操作的结果也给出了所有的参数显著性检验。以下以 Stata 13 为例，采用最大似然估计法ＭＬＥ，得到以下结果：

ARCH family regression -- AR disturbances

Sample: 1947:2 - 2012:1	Number of obs = 260
Distribution: Gaussian	Wald chi2(1) = 38.91
Log likelihood = 866.4794	Prob > chi2 = 0.0000

growth	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
growth						
_cons	.0086687	.0008492	10.21	0.000	.0070042	.0103332
ARMA						
ar						
L1.	.4383648	.0702734	6.24	0.000	.3006316	.5760981
ARCH						
arch						
L1.	.330808	.083904	3.94	0.000	.1663592	.4952569
L2.	.3870111	.1109962	3.49	0.000	.1694626	.6045595
_cons	.0000344	5.33e-06	6.46	0.000	.000024	.0000449

从中可以看到，整个模型的表达式为：

$$\begin{cases} growth_t = 0.0097 + 0.4384 growth_{t-1} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t,\ \ e_t \sim N(0,1) \\ h_t = 0.0000344 + 0.3308\varepsilon_{t-1}^2 + 0.3870\varepsilon_{t-2}^2 \end{cases} \tag{9.19}$$

同时，所有参数的 p-value 值都小于 0.05 ，即通过了参数显著性检验。

Step 8：**模型显著性检验。**

Step 9：**分布检验。**