第九讲：非平稳序列：异方差模型（一）

概要： 本讲介绍方差非平稳序列分析。主要介绍异方差问题的产生、方差齐性变化原理，以及 ARCH/GARCH 模型的产生机理、模型结构及一些基础的衍生模型。

一、异方差问题

20 世纪 50 年代开始，金融界开始证券投资理论 Portfolio Theory 的研究，波动 Volatility 成为一个极其重要的概念，经常出现在资产定价模型 Assent Pricing 和风险管理 Risk Management 等模型中。波动有各种各样的定义，在时间序列分析中，波动被认为是**一段变化且与高可变性或高方差相关的时间序列**。研究者观察到，时间序列具有交替时期变化的特征，在相对平静时期，波动性较低，而相对变动时期，波动性相对比较高，即具有集群效应。

异方差产生的原因是时序数据存在集群效应（volatility cluster），它是指在消除确定性非平稳因素的影响后，残差序列在大部分时段小幅波动，但是会在某些时段出现持续大幅波动，于是序列的波动就呈现出一段持续时间的小幅波动和一段持续时间的大幅波动交替出现的特征。**集群效应的产生原因，通常认为是经济市场和金融市场的波动易受谣言、政局变动、政府货币与财政政策变化等诸多因素的影响**，一旦某个影响因素出现，市场会大幅波动，以消化这个影响，这就出现密集的大幅波动，波动到位实现新的稳定之后，在下一个影响因素到来之前，序列会维持一段时间的小幅波动。典型的集群效应图见图 8.1。

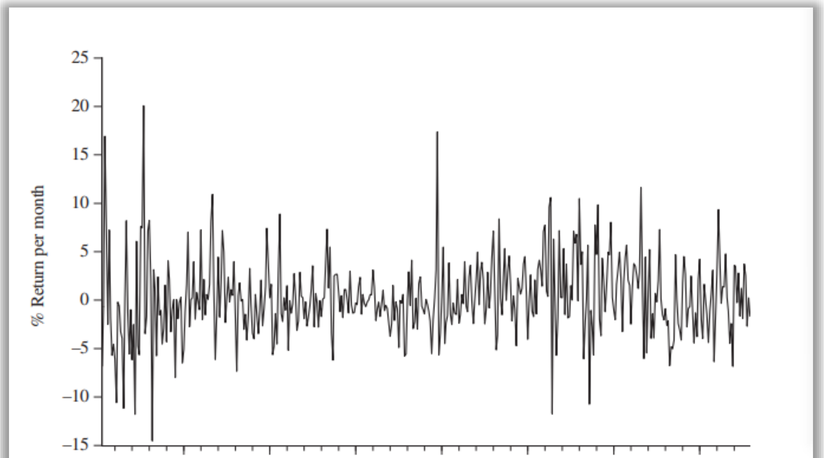

图 8.1 1980 年-2017 年黄金价格月度回报率

对波动性研究兴趣很大程度上与它不能直接观察有关。20 世纪 80 年代初，有人提出波动性应该嵌入观测到的时间序列的正式随机模型中，即，尽管一些序列似乎是序列不相关的，或者说均值之间是无关的，或者说一阶矩 first moment 和时间无关，但并不能断定此序列的所有统计特征都和时间 t 无关。这些时序有可能较高的“矩 Moment”表现出丰富的动态性（即和时间 t 有关），也就是说，时序分析不仅要关注均值（一阶矩）的建模，还应该注意序列的高“矩 Moment”特征的建模。

时间序列的宽平稳牵涉到的统计特征是**一阶矩和二阶矩**，一阶矩就是随机变量的期望（均值），二阶矩就是随机变量的方差及协方差。要构建稳定的关系，协方差必须只与时间间隔 lag 相关，而与所处的时间 t 无关。因此，研究高矩的变动，只能聚焦于随机变量的方差。虽然平稳过程必须有恒定的方差，但某些条件方差可能会改变，因此尽管无条件方差 $Var(X_t)$ 可能对所有 t 都是恒定的，但条件方差 $Var(X_t|X_{t-1,}X_{t-2,}\dots,X_{t-k,})$，取决于 X_t 的实现，即一方面取决于 k 的大小，另一方面，随着时间推移，$X_{t-1,}X_{t-2,}\dots,X_{t-k,}$ 对应的观测值 $x_{t-1,}x_{t-2,}\dots,x_{t-k,}$ 也在变化。

具有随时间变化方差的随机过程可以看作由以下乘积过程 Product Process 产生，即：

$$X_t = \mu + \sigma_t e_t \qquad 8.1$$

其中，e_t 是标准正态分布，即 $E(e_t) = 0$，$Var(e_t) = E(e_t^2) = 1$，σ_t 是一正的随机变量，则：

$$Var(X_t|\sigma_t) = E[(X_t - \mu)^2|\sigma_t] = \sigma_t^2 E(e_t^2) = \sigma_t^2 \qquad 8.2$$

σ_t^2 就是条件方差 Conditional Variance，而 σ_t 是条件标准差 Conditional Standard Deviation 。

$$e_t = \frac{X_t - \mu}{\sigma_t} \qquad 8.3$$

式 8.3 为标准正态分布，是一纯随机过程（白噪声），即 $E(e_t e_{t-k}) = 0$，$k \neq 0$，则 $\{X_t\}$ 的均值为 μ，方差为：

$$E(X_t - \mu)^2 = E(\sigma_t^2 e_t^2) = E(\sigma_t^2)E(e_t^2) = E(\sigma_t^2) \qquad 8.4$$

自协方差为：

$$E(X_t - \mu)(X_{t-k} - \mu) = E(\sigma_t \sigma_{t-k} e_t e_{t-k}) = E(\sigma_t \sigma_{t-k})E(e_t e_{t-k}) = 0 \qquad 8.5$$

二、方差齐性变化

如果存在异方差，最为直接的方法就是将原时序通过一定的函数变换，生成一个新的时序，而这个时序的方差变动将大大减少，或者方差不变了（称为方差齐性）。这里的关键是要找到合适的转换函数。在《计量经济学》中提到，最为常用的转换函数是取对数。

通常$\{X_t\}$方差σ_t^2和均值μ_t之间存在着某种函数关系，即：

$$Var(X_t) = \sigma_t^2 = f(\mu_t) \qquad 8.6$$

其中，$f(\ )$是某个已知的函数。记$g(\ )$为转换函数，使得转换后的$g(X_t)$满足方差齐性：

$$Var[g(X_t)] = \sigma^2 \qquad 8.7$$

将$g(X_t)$在μ_t附近做一阶泰勒展开：

$$g(X_t) \approx g(\mu_t) + (X_t - \mu_t)g'(\mu_t) \qquad 8.8$$

其中，$g'(\ )$是$g(\ )$的一阶导数。则$g(X_t)$的方差近似等于：

$$Var[g(X_t)] \approx Var[g(\mu_t) + (X_t - \mu_t)g'(\mu_t)] = [g'(\mu_t)]^2 Var(X_t)$$

$$= [g'(\mu_t)]^2 f(\mu_t) \qquad 8.9$$

要使得$Var[g(X_t)]$等于常数，转换函数$g'(\ )$和$\sqrt{f(\ )}$必须具有倒数关系，即：

$$g'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}} \qquad 8.10$$

在现实中，许多金融序列都有异方差，而且序列的标准差和水平均值之间具有某种正比关系，即序列的水平低时，序列的波动方位小，序列水平高时，序列的波动范围大。对于这种异方差的性质，通常假定：

$$\sigma_t = \mu_t \qquad 8.11$$

等价于：

$$\sigma_t^2 = f(\mu_t) = \mu_t^2 \qquad 8.12$$

要是的原序列经过转换后方差齐性，就必须满足：

$$g'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}} = \frac{1}{\mu_t} \qquad 8.13$$

积分得到；

$$g(\mu_t) = ln(\mu_t) \qquad 8.14$$

即对于标准差和水平成正比关系的异方差序列，对数变换可以有效实现方差齐性。

三、ARCH/GARCH 模型的产生机理

集群效应意味着序列的波动存在相关性。因为如果序列的波动不存在相关性的话，就不会产生小幅波动和大幅波动集中交替出现，而是呈现出大幅波动和小幅波动完全无规律。基于这个思想，Engle 构造了自回归条件异方差模型（Auto-Regressive Conditional Heteroskedastic Model）

2.1 理论阐述

设定条件方差σ_t^2与历史信息$X_{t-1}, X_{t-2}, \ldots$有关，写成函数形式为：

$$\sigma_t^2 = f(X_{t-1}, X_{t-2}, \ldots) \qquad 8.15$$

最为简单的模型是只与X_{t-1}有关，依据式 8.1 的构建思想，可以写出最简单的模型：

$$\sigma_t^2 = f(X_{t-1}) = b_0 + b_1(X_{t-1} - \mu)^2 \qquad 8.16$$

其中，b_0, b_1都取正值，以保证$\sigma_t^2 > 0$，$(X_{t-1} - \mu)$度量了X_{t-1}偏离中心值μ的程度，即离差，为保证方差为正，所以要平方，式 8.16 的含义就是方差和离差平方呈线性关系。在$e_t \sim N(0,1)$的前提下，依据式$X_t = \mu + \sigma_t e_t$可知，X_t呈现条件正态分布，即：

$$X_t | X_{t-1}, X_{t-2}, \ldots \sim N(\mu, \sigma_t^2)$$

因此，

$$Var(X_t | X_{t-1}) = \sigma_t^2 = b_0 + b_1(X_{t-1} - \mu)^2 \qquad 8.17$$

记：$\varepsilon_{t-1} = X_{t-1} - \mu = \sigma_t e_t$，$\varepsilon_t$是残差，则

$$\varepsilon_t | X_{t-1}, X_{t-2}, \ldots \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = b_0 + b_1 \varepsilon_{t-1}^2 \qquad 8.18$$

式 8.18 就是 1 阶自回归条件异方差模型 first-order autoregressive conditional heteroskedastic ARCH(1)，注意：教材中h_t就是这里的σ_t^2。

定义：$\alpha_t = \varepsilon_t^2 - \sigma_t^2$，则式 8.9 就可以写成：

$$\varepsilon_t^2 = b_0 + b_1 \varepsilon_{t-1}^2 + \alpha_t \qquad 8.19$$

由于$E(\alpha_t | X_{t-1}, X_{t-2}, \ldots) = 0$，式 8.10 可以看成是针对残差序列$\{\varepsilon_t^2\}$的 AR(1)模型，而

$$\alpha_t = \varepsilon_t^2 - \sigma_t^2 = \sigma_t^2(e_t^2 - 1) \qquad 8.20$$

说明α_t存在着随时间变化的方差。

将式 8.17 推广到X_{t-q}就得到 ARCH(q)：

$$\sigma_t^2 = f\left(X_{t-1}, X_{t-2}, \ldots, X_{t-q}\right) = b_0 + b_1(X_{t-1} - \mu)^2 + \cdots + b_q\left(X_{t-q} - \mu\right)^2 \qquad 8.21$$

其中，$b_0, b_1 \ldots b_q$都是正值，令$\varepsilon_{t-1} = X_{t-1} - \mu$，$\varepsilon_{t-2} = X_{t-2} - \mu, \ldots$，$\varepsilon_{t-q} = X_{t-q} - \mu$，则式 8.19 可以写成：

$$\sigma_t^2 = b_0 + b_1 \varepsilon_{t-1}^2 + \cdots + b_q \varepsilon_{t-q}^2 \qquad 8.22$$

同样，$\alpha_t = \varepsilon_t^2 - \sigma_t^2$，则：

$$\varepsilon_t^2 = b_0 + b_1 \varepsilon_{t-1}^2 + \cdots + b_q \varepsilon_{t-q}^2 + \alpha_t \qquad 8.23$$

ARCH(q)模型的一个实际困难是，当 q 很大时，无约束估计通常会导致某些b_i的值小

于零，这样就无法确保条件方差σ_t^2总是正的。在早期应用中，为避免这种情况出现，只能限定 q 取一个很小的值。为了获得更大的灵活性，引入了 generalized autoregressive conditional heteroskedastic GARCH，也称为广义自回归条件异方差模型，GARCH(p,q)具有如下模型结构：

$$\sigma_t^2 = b_0 + b_1\varepsilon_{t-1}^2 + \cdots + b_q\varepsilon_{t-q}^2 + c_1\sigma_{t-1}^2 + c_2\sigma_{t-2}^2 + \cdots + c_p\sigma_{t-p}^2 \qquad 8.24$$

其中，$p > 0$，$c_i \geq 0$。

2.2 直观理解

对于时间序列X_1 X_2 $\ldots X_t$存在异方差，说明每个时刻对应变量的方差是不同的。即σ_1^2 σ_2^2 $\ldots\sigma_t^2$是不同的，即方差也构成了一个时间序列。如果这些方差的值是随机产生的，它们之间不存在着任何关系，即σ_1^2 $\sigma_2^2 \ldots\sigma_t^2$是白噪声，那么对此方差时间序列无法进行深入研究；如果他们之间存在着相关关系，那么就要把这种数量变化关系用模型的形式表达出来，表达的基本思路就是回归分析。

我们现有的只有一个观测值序列x_1 x_2 $\ldots x_t$，它的方差大小$\widehat{\sigma_t}^2$是由$\widehat{\varepsilon_1}$ $\widehat{\varepsilon_2}$ $\ldots\widehat{\varepsilon_t}$的方差决定的。$\widehat{\sigma_t}^2 = Var(\widehat{\varepsilon_t}) = E(\widehat{\varepsilon_t}^2)$。现在我们能观察到的$\widehat{\varepsilon_1}^2$ $\widehat{\varepsilon_2}^2 \ldots\widehat{\varepsilon_t}^2$，是一个时间序列，是回归残差平方值所构成的时间序列。如果他们之间存在着滞后 p 期的关系，即 AR(p)模型：

$$\widehat{\varepsilon_t}^2 = \lambda_0 + \lambda_1\widehat{\varepsilon_{t-1}}^2 + \lambda_2\widehat{\varepsilon_{t-2}}^2 + \ldots + \lambda_p\widehat{\varepsilon_{t-p}}^2 + \alpha_t \qquad 8.25$$

这里的α_t是残差平方序列$\{\widehat{\varepsilon_t}^2\}$AR(p)模型的残差项，对于观测值序列$\widehat{\varepsilon_1}^2$ $\widehat{\varepsilon_2}^2 \ldots\widehat{\varepsilon_t}^2$来说

$$\alpha_t = \widehat{\varepsilon_t}^2 - E\left(\widehat{\varepsilon_t}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p}}^2\right) \qquad 8.26$$

最后式子变成：

$$\widehat{\varepsilon_t}^2 = \lambda_0 + \lambda_1\widehat{\varepsilon_{t-1}}^2 + \lambda_2\widehat{\varepsilon_{t-2}}^2 + \ldots + \lambda_p\widehat{\varepsilon_{t-p}}^2 + \widehat{\varepsilon_t}^2 - E\left(\widehat{\varepsilon_t}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p}}^2\right) \qquad 8.27$$

$$E\left(\widehat{\varepsilon_t}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p}}^2\right) = \omega + \lambda_1\widehat{\varepsilon_{t-1}}^2 + \lambda_2\widehat{\varepsilon_{t-2}}^2 + \ldots + \lambda_p\widehat{\varepsilon_{t-p}}^2 \qquad 8.28$$

这就是 ARCH(p)模型，可以视为就是$\widehat{\varepsilon_1}^2$ $\widehat{\varepsilon_2}^2 \ldots\widehat{\varepsilon_t}^2$的 AR(p)模型的变换。注意：这里的$E(\widehat{\varepsilon_t}^2 | \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p}}^2)$就是教材中的$h_t$，上一节中的$\sigma_t^2$

如果$\widehat{\varepsilon_1}^2$ $\widehat{\varepsilon_2}^2 \ldots\widehat{\varepsilon_t}^2$的关系符合 ARMA($p$,$q$)模型，即：

$$\widehat{\varepsilon_t}^2 = \lambda_0 + \lambda_1\widehat{\varepsilon_{t-1}}^2 + \lambda_2\widehat{\varepsilon_{t-2}}^2 + \ldots + \lambda_p\widehat{\varepsilon_{t-p}}^2 + \alpha_t + \delta_1\alpha_{t-1} + \delta_2\alpha_{t-2} + \ldots \delta_q\alpha_{t-q} \qquad 8.29$$

由于 $\alpha_t = \widehat{\varepsilon_t}^2 - E(\widehat{\varepsilon_t}^2 | \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p}}^2)$，

$$\alpha_{t-1} = \widehat{\varepsilon_{t-1}}^2 - E(\widehat{\varepsilon_{t-1}}^2 | \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p-1}}^2)$$

....

$$\alpha_{t-q} = \widehat{\varepsilon_{t-q}}^2 - E\left(\widehat{\varepsilon_{t-q}}^2 \middle| \widehat{\varepsilon_{t-q-1}}^2, \ldots \widehat{\varepsilon_{t-p-q}}^2\right) \qquad 8.30$$

则得到：

$$\widehat{\varepsilon_t}^2 = \omega + \lambda_1 \widehat{\varepsilon_{t-1}}^2 + \lambda_2 \widehat{\varepsilon_{t-2}}^2 + \cdots + \lambda_p \widehat{\varepsilon_{t-p}}^2 + \widehat{\varepsilon_t}^2 - E\left(\widehat{\varepsilon_t}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \dots \widehat{\varepsilon_{t-p}}^2\right)$$

$$+ \delta_1 \left(\widehat{\varepsilon_{t-1}}^2 - E\left(\widehat{\varepsilon_{t-1}}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \dots \widehat{\varepsilon_{t-p-1}}^2\right)\right) + \ldots + \delta_q \left(\widehat{\varepsilon_{t-q}}^2\right.$$

$$\left. - E\left(\widehat{\varepsilon_{t-q}}^2 \middle| \widehat{\varepsilon_{t-q-1}}^2, \ldots \widehat{\varepsilon_{t-p-q}}^2\right)\right) \qquad 8.31$$

移项得

$$E\left(\widehat{\varepsilon_t}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \dots \widehat{\varepsilon_{t-p}}^2\right) = \omega + (\lambda_1 + 1)\widehat{\varepsilon_{t-1}}^2 + (\lambda_2 + 1)\widehat{\varepsilon_{t-2}}^2 + \cdots$$

$$+ (\lambda_p + 1)\widehat{\varepsilon_{t-p}}^2 + \delta_1 E\left(\widehat{\varepsilon_{t-1}}^2 \middle| \widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p-1}}^2\right) \ldots + \delta_q \left(\widehat{\varepsilon_{t-q}}^2\right.$$

$$\left. - E\left(\widehat{\varepsilon_{t-q}}^2 \middle| \widehat{\varepsilon_{t-q-1}}^2, \ldots \widehat{\varepsilon_{t-p-q}}^2\right)\right) \qquad 8.32$$

这就是 GARCH(q,p)模型，可以视为是$\widehat{\varepsilon_1}^2$ $\widehat{\varepsilon_2}^2$ ...$\widehat{\varepsilon_t}^2$的 ARMA($p$,$q$)模型的转换。注意：这里的$E(\widehat{\varepsilon_t}^2|\widehat{\varepsilon_{t-1}}^2, \ldots \widehat{\varepsilon_{t-p}}^2)$就是教材中的$h_t$，上一节中的$\sigma_t^2$。

显然，$\widehat{\varepsilon_1}^2$ $\widehat{\varepsilon_2}^2$ ...$\widehat{\varepsilon_t}^2$和其他时序不同的地方就是多了一个限制条件：$\widehat{\varepsilon_t}^2 \geq 0$。

总之，ARMA/ARIMA 模型估计的是X_1 X_2 ...X_t，每一个X的均值，即点估计。而 ARCH/GARCH 估计的是每一个X的变动范围。二者结合，可以说得到了$\{X_t\}$的区间估计。

四、ARCH/GARCH 模型结构

假定在历史数据已知的情况下，零均值残差序列具有异方差性，即：

$$E(\varepsilon_t) = 0 \qquad Var(\varepsilon_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = h_t$$

而且残差平方序列$\{\varepsilon_t^2\}$具有相关性，则 ARCH(q)模型的结构式如下：

$$h_t = \lambda_0 + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2 \qquad 8.33$$

式 8.33 需要满足很强的参数约束条件：

（1）方差必须非负。即式 8.33 必须大于或等于零，要保证对任意取值的ε_{t-j}都成立，只能要求每个参数都非负，即：$\lambda_0, \lambda_1, \dots, \lambda_q \geq 0$；

（2）序列无条件方差存在，即$Var(\varepsilon_t) = \sigma^2$，则对式 8.33 求期望，得：

$$E[Var(\varepsilon_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)] = E[\lambda_0 + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2]$$

$$E[E(\varepsilon_t^2|\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)] = \lambda_0 + \sum_{i=1}^{q} \lambda_i E(\varepsilon_{t-i}^2)$$

$$\sigma^2 = \lambda_0 + \sum_{i=1}^{q} \lambda_i \sigma^2$$

$$\sigma^2 = \frac{\lambda_0}{1 - \sum_{i=1}^{q} \lambda_i}$$

要保证无条件方差σ^2存在且方差不能为负，分母要大于零，即：

$$1 - \sum_{i=1}^{q} \lambda_i > 0$$

综合（1）和（2），ARCH(q)模型参数要满足以下条件：

$$\lambda_0 > 0 \qquad 0 \le \lambda_i < 1 \quad i = 1,2,\dots,q \qquad 且\, \lambda_1 + \lambda_2 + \cdots + \lambda_q < 1$$

ARCH 模型的实质是使用残差平方序列的 q 阶移动平均拟合当期异方差函数值。由于移动平均模型具有自相关系数 q 阶截尾，所以 ARCH 模型只适用于异方差函数短期自相关过程。但在实际中，有些异方差函数具有长期的自相关性，此时采用 ARCH 模型，将会产生很高的移动平均阶数，这会影响 ARCH 模型的拟合精度。Engle 的学生 Tim Bollerslev 对 ARCH 模型进行改进，引入了异方差的历史信息h_{t-j}，构建了 GARCH 模型，其提取的波动信息更加充分。

GARCH(p,q)模型的结构式如下：

$$h_t = \lambda_0 + \sum_{j=1}^{p} \eta_j h_{t-j} + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2 \qquad 8.34$$

参考式 8.33 约束条件（1）和（2）的推导，得到式 8.34 的约束条件如下：

$$\lambda_0 > 0$$

$$0 \le \lambda_i < 1 \qquad i = 1,2,\dots,q$$

$$0 \le \eta_j < 1 \qquad j = 1,2,\dots,q$$

$$0 \le \sum_{j=1}^{p} \eta_j + \sum_{i=1}^{q} \lambda_i < 1$$

为了更清楚地理解 GARCH 模型的实质，把异方差序列$\{h_t\}$视为响应序列，把残差平方序列$\{\varepsilon_t^2\}$视为随机扰动项序列，那么，ARCH 模型实际上就是$\{h_t\}$关于$\varepsilon_t^2, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2 \dots \varepsilon_{t-q}^2$的 q 阶移动平均模型 MA(q)；GARCH 模型就是$\{h_t\}$关于$h_{t-1}, h_{t-2}, \dots, h_{t-p}$的 p 阶自相关，关于$\varepsilon_t^2, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2 \dots \varepsilon_{t-q}^2$的 q 阶移动平均的 ARMA(p,q)模型。ARCH 模型是 GARCH 模型的一个特例。

五、GARCH 衍生模型

GARCH 模型给出了对波动性进行描述的方法，为大量的金融时间序列提供了有效的分析工具，它是迄今为止最常用、最便捷的异方差序列拟合模型。但大量的使用经验表明，它也存在着一些不足。

其一，是对参数的约束非常严格。方差非负的要求导致参数非负的约束条件，即：

$$\lambda_0 > 0 \quad \lambda_i \geq 0 \quad \eta_j \geq 0$$

同时，无条件方差必须平稳，要求参数有界：

$$\sum_{j=1}^{p} \eta_j + \sum_{i=1}^{q} \lambda_i < 1$$

这些约束条件在一定程度上限制了 GARCH 模型的适用范围。

其二，对正负扰动的反应是对称的。扰动项是真实值和预测值之差，如果扰动项为正，说明真实值比预测值大，对于投资者而言就是获得超预期收益；如果扰动项为负，说明真实值比预测值小，对于投资者而言就是出现了超预期亏损。以 ARCH(1)模型为例（采用残差平方的 AR 模型表示）：

$$\widehat{\varepsilon_t}^2 = \lambda_0 + \lambda_1 \widehat{\varepsilon_{t-1}}^2$$

即无论$\widehat{\varepsilon_{t-1}}$是正是负，它对下一期的影响都是$\lambda_1$，这意味着无论上一期的投资是盈利还是亏损，对投资人下一期投资行为的影响都是一样的，这与现实情况不符。大量实践表明，投资人在出现收益时，通常反应比较慢，出现亏损时，通常反应比较块，这种情况，如果不在模型中考虑，会影响预测精度。

为了拓宽 GARCH 模型的适用范围，提高模型的拟合精度，学者从不同角度出发，构造了许多 GARCH 模型的衍生模型。以下主要介绍三个衍生模型。

5.1 EGACH 指数 GARCH 模型

Nelson 于 1991 年提出了 EGARCH 模型，其具体结构如下：

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \dots) + \varepsilon_t \\ \varepsilon_t = e_t \sqrt{h_t} \\ lnh_t = \lambda_0 + \sum_{j=1}^{p} \eta_j lnh_{t-j} + \sum_{i=1}^{q} \lambda_i g(e_{t-j}) \\ g(e_t) = \theta e_t + \gamma(|e_t| - E|e_t|) \end{cases} \qquad 8.35$$

式 8.35 中，第一个式子是拟合均值的模型；第二个式子是残差分布的假设条件，其中：e_t是标准正态分布；第三个式子是拟合方差的模型，这里lnh_t是条件方差的对数，它可以是

正数，也可以是负数，这说明，第三个式子的系数参数不需要任何非负假定，EGARCH 模型的第一个改进是放松了 GARCH 模型的参数约束。

EGARCH 模型的第二个改进是引入了加权扰动函数$g(e_t)$，通过特殊的函数构造，能对正负扰动进行非对称处理。

$$g(e_t)=\theta e_t+\gamma(|e_t|-E|e_t|)=\begin{cases}(\theta+\gamma)e_t-\gamma E|e_t|, & e_t>0\\(\theta-\gamma)e_t-\gamma E|e_t|, & e_t<0\end{cases} \qquad 8.36$$

其中，e_t是标准正态分布，$E|e_t|=\sqrt{2/\pi}$，通常$\gamma=1$。

5.2 IGARCH　方差无穷 GARCH 模型

当 GARCH 模型平稳时，ε_t的无条件方差为：

$$Var(\varepsilon_t)=\frac{\lambda_0}{1-\left(\sum_{j=1}^{p}\eta_j+\sum_{i=1}^{q}\lambda_i\right)} \qquad 8.37$$

参数的约束条件保证了$Var(\varepsilon_t)$有界，GARCH 模型能实现宽平稳。如果把约束条件改为：

$$\sum_{j=1}^{p}\eta_j+\sum_{i=1}^{q}\lambda_i=1 \qquad 8.38$$

就构成了 IGARCH 模型。此时，$Var(\varepsilon_t)$无界，所以它的无条件方差无意义。

IGARCH 模型适用描述具有单位根特征（随机游走）的条件异方差。从理论角度分析，IGARCH 现象可能是由波动率带有常数飘移项引起的。

5.3 GARCH-M　依均值 GARCH 模型

在金融领域，风险厌恶型投资者会要求资产的收益率和波动性相匹配，，即：

风险投资期望收益=无风险收益+风险溢价　　8.39

式 8.39 中，无风险收益为资金的时间价值，是投资者从事风险极小的投资(诸如国债、货币市场工具或银行存款)所获得的收益。超额收益是任何特定时期风险资产同无风险资产收益之差。风险溢价是超额收益的期望。风险溢价指投资者因承担风险而获得的额外报酬。风险越大，风险溢价越高。

Engle, Lilien 和 Robins(1987)将风险溢价思想引入 GARCH 模型，允许序列的均值依赖于它的波动性，由此提出 GARCH-M 模型。

GARCH-M 模型的构造思想是：序列均值与条件方差之间具有某种相关关系，这时可以把条件标准差作为附加回归因子建模，模型结构如下：

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \ldots) + \delta\sqrt{h_t} + \varepsilon_t \\ \varepsilon_t = e_t\sqrt{h_t} \\ h_t = \lambda_0 + \sum_{j=1}^{p} \eta_j h_{t-j} + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2 \end{cases} \tag{8.40}$$