第 7 章 波动率的 GARCH 模型*

F.C.Palm

1．引言

直到大约 15 年前，有关时间序列统计分析的焦点才集中在条件一阶矩。风险和不确定性在经济决策模型中发挥着越来越重要的作用，衡量风险和波动率的一般方法显现出明显的随时间变动的特性，因而导致对二阶矩中的时间变差（Time-Variation）建模的时间序列方法的新发展。

Engle(1982)提出了条件方差的自回归条件异方差模型（ARCH），它与 Box-Jenkins 类型的条件一阶矩模型一致，在分析经济时间序列时非常有用。从那以后，很多文献提出对更高阶条件矩的模型化。在金融时间序列领域可以见到它们的许多应用。Bollerslev 等（1992 和 1994，Nijman 和 Palm（1993），Diebold 和 Lopez（1994），Pagan（1995）以及 Bera 和 Higgings（1995）都对 ARCH 模型做了广泛的理论和实证研究。Gourieroux（1992）还更具体地研究了 ARCH 模型，使它更加合乎规范。

这一章的目的是对金融领域中使用 ARCH 和 GARCH（广义 ARCH）进行条件波动率建模的某些方面作选择性说明，并将 ARCH 方法与其他可供选择的研究思路相比较。其重点是介绍近期成果，例如利用因子－ARCH 建立多变量模型，最后还将讨论模型评价。

在第 2 节，我们将引入单变量和多变量 GARCH 模型（包括 ARCH 模型），讨论它们的特点和函数形式的选择，将它们与其他的波动率模型相比较。第 3 节讨论这些模型的推断问题。第 4 节讨论 GARCH 模型的统计特征，它们与连续时间扩散（diffusion）模型的关系以及波动率预测。最后，第 5 节进行总结并评论潜在的未来研究方向。

2．GARCH 模型

2.1 动机

GARCH 模型的提出是为了解释金融数据的经验规律。正如 Pagan（1995）、Bollerslev 等（1994）强调的那样，许多金融时间序列具有共同特征。第一，资产价格通常是非平稳的，有单位根，而收益率通常是平稳的。越来越多的证据证明一些金融序列是分数单整（fractionally integrated）的。第二，收益率序列通常表现出无自相关或很小的自相关。平方值序列的相互独立经常被拒绝，表明子序列的观测值之间存在非线性关系。收益序列的波动呈现“群集”性，大的波动发生在更长的时期，低收益往往跟着小波动值。这些现象都与依时间变化的条件变量有关。第三，由于接受某些肥尾分布，正态性往往被拒绝。序列中存在着的无条件超峰度可能与条件方差中的时间变差有关。第四，一些序列显示出所谓的杠杆效应[参见 Black （1976）]，即股价的变动倾向于与波动率的变化呈负相关。一些序列具有偏斜的无条件经验分布表明正态分布是不合适的。第五，不同证券的波动率经常一起变动表明市场与某些共同因子之间的联系可以解释条件二阶矩中的时频变差（temporal variatioin）。在下一节，我们将提出一些模型，来解释条件方差之间的时频依赖关系、偏度及超峰度。

---

* 作者对 G.S.Maddala 对本文早期版本的许多有益的评论表示感谢。

2.2 单变量 GARCH 模型

考虑下面形式的随机模型

$$y_t = \varepsilon_t h_t^{1/2} \tag{2.1}$$

$$h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 \tag{2.2}$$

其中 $E\varepsilon_t = 0, Var(\varepsilon_t) = 1, \alpha_0 > 0, \beta_i \ge 0, \alpha_i \ge 0,$ 且 $\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i < 1$。这是一个 (p,q) 阶的 GARCH 模型，由 Bollerslev（1986）提出。当 $\beta_i = 0, i = 1,2,\cdots,p$ 的时候，它就是特殊的 ARCH（q）模型，由 Engle（1982）在一篇开创性的论文中提出来。非负条件隐含着非负方差，但广义平稳性要求对 α_i 与 β_i 之和予以限制。正如 Nelson 和 Cao（1992）指出的那样，非负条件方差的充分条件很大程度上被弱化了。如果 y_t^2 过去的实现值大于 σ^2， y_t 的条件方差将大于由 $\sigma^2 = \alpha_0 / (1 - \sum_{i=1}^{p} \beta_i - \sum_{i=1}^{q} \alpha_i)$ 给出的无条件方差。Anderson（1992）已证明，GARCH 模型属于确定性条件异方差类的模型，此类模型中，条件方差是一个 t 时可得信息集变量的函数。在正态假设下，模型可写为：

$$y_t \mid \Phi_{t-1} \sim N(0, h_t), \tag{2.3}$$

其中 h_t 由（2.2）给出，Φ_{t-1} 代表在 t-1 时的信息。Anderson（1994）对确定型、条件异方差型、条件随机型和同期随机型的波动过程作了区分。不那么严格地说，如果直到 t=0 时（包括 t＝0），信息集（σ—域）Φ 等于系统中所有随机向量的 σ －域，那么波动过程是确定性的；如果 Φ 包括信息变量并且在 t-1 时可观测，那么波动过程是条件异方差性的；如果 Φ 包括直到 t-1 时的全部随机向量，那么波动过程是条件随机性的；如果 Φ 的信息包括直到 t 时的全部随机向量，那么波动过程是同期随机性的。这里要注意的是施加于各种波动率表述形式的信息结构的顺序。

当 $\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i = 1$，就产生了单整的 GARCH（integrated GARCH， IGARCH）模型[参见 Engle 和 Bollerslev (1986)]。从 (2.2) 的 GARCH(p,q) 模型，可得到 $[1 - \alpha(L) - \beta(L)] y_t^2 = \alpha_0 + [1 - \beta(L)] v_t$，其中 $v_t = y_t^2 - h_t$ 是条件方差过程的新生(innovation)，$\alpha(L) = \sum_i^q \alpha_i L^i$，$\beta(L) = \sum_{i=1}^{p} \beta_i L^i$。如果以滞后算子 L 表达的多项式 $1 - \alpha(L) - \beta(L)$ 可以分解为 $\Phi(L)(1-L)^d$，其中 $\Phi(z) = 0$ 的根在单位圆外，且 $0 \le d \le 1$，就产生了分数单整的 GARCH 模型(FIGARCH（p,d,q）)，它是由 Ballie、Bollerslev 和 Mikkelsen（1993）提出的。当 d=0，FIGARCH 模型嵌套 GARCH（p,q）模型； 当 d=1，它嵌套 IGARCH（p,q）模型。允许 d 在 0 到 1 之间取值，就有更大的灵活性，这在建立条件方差的长期依赖关系模型时可能相当重要。

在对金融数据作实证分析时，发现 GARCH（1,1）或 GARCH（1,2）通常可以适当地解释条件异方差性。这个发现类似于低阶 ARMA 模型总能很好地描述许多经济时间序列的条件均值的动态变化。

在以上模型中，正的过去值与负的过去值对条件方差的效应是对称的，认识到这一点非常重要。然而许多金融序列却明显不对称。负的股票收益引起波动率增大的幅度要大于同样的正收益的效应。Black（1976）把这一现象解释为杠杆效应的结果：股票价值的大幅度下降并未使债务价值相应地下降，因而提高了债务－权益比率(debt-equity ratio)。诸如 Nelson（1991）提出的指数 GARCH（exponential GARCH，EGARCH）模型、Sentana（1991）和 Engle（1990）提出的二次 GARCH（quadratic GARCH，QGARCH）模型、Zakoian（1994）提出的阈 GARCH（threshold TGARCH）模型都允许存在不对称。

Nelson 的 EGARCH 模型如下：

$$\ln h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i \ln h_{t-i} + \sum_{i=1}^{q} \alpha_i (\varphi \varepsilon_{t-i} + \phi \,|\, \varepsilon_{t-i} \,| - \phi E \,|\, \varepsilon_{t-i} \,|) \quad , \qquad (2.4)$$

其中参数 α_0、α_i、β_i 不受非负限制。当 $\alpha_i > 0$ 且 $\varphi < 0$*时，可以解释为收益的负冲击提高了债务－股东权益比率，从而增加了未来收益的不确定性。类似地，在指数模型中考虑到分数单整时，就得到了 FIEGARCH 模型。

Sentana（1991）把 QGARCH 模型写成

$$h_t = \sigma^2 + \psi' x_{t-q} + x'_{t-q} A x_{t-q} + \sum_{i=1}^{p} \beta_i h_{t-i} \quad , \qquad (2.5)$$

其中 $x_{t-q} = (y_{t-1}, y_{t-2}, \cdots, y_{t-q})'$。该线性形式考虑到了非对称性。矩阵 A 非对角线上的元素解释了 x_t 的滞后值与条件方差相互作用的结果。（2.5）中嵌入了文献中所提出的各种二次方差函数。Bera 和 Lee（1990）假设 $\psi = 0$，提出了增广的 GARCH 模型（GGARCH）。Engle（1982）提出的 ARCH 模型中，约束 $\psi = 0$、$\beta_i = 0$，且 A 为对角阵。Engle（1990）以及 Engle 与 Ng（1993）提出的非对称 GARCH 模型，也假设 A 为对角阵。Robinson（1991）研究了线性标准差模型，约束 $\beta_i = 0$、$\sigma^2 = p^2$、$\psi = 2\rho\varphi$，而且 $A = \varphi\varphi'$ 是秩为 1 的矩阵。这样，条件方差就变成 $h_t = (\rho + \varphi' x_{t-q})^2$。

Zakoian（1994）提出的 TGARCH 模型由下式给出：

$$h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} (\alpha_i^+ y_{t-i}^+ + \alpha_i^- y_{t-i}^-) \quad , \qquad (2.6)$$

其中 $y_t^+ = \max\{y_t, 0\}$，$y_t^- = \min\{y_t, 0\}$。它通过参数 α_i^+ 和 α_i^- 的不同来解释非对称性。

正如 Hentschel（1994）所证明的那样，许多 GARCH 类模型（取 p=q=1）能够包含在绝对 GARCH（Absolute GARCH，AGARCH）模型的 Box-Cox 变换中

$$(\sigma_t^{\lambda} - 1)/\lambda = \alpha_0 + \alpha_1 \sigma_{t-1}^{\lambda} f^{\nu}(\varepsilon_{t-1}) + \beta(\sigma_{t-1}^{\lambda} - 1)/\lambda \quad , \qquad (2.7)$$

* 原书为 $\alpha_i \succ 0$ 且 $\varphi \prec 0$，有误。下页的 $\lambda \succ 1$ 和 $\lambda \prec 1$ 情况相同。——译者注。

其中$\sigma_t = h_t^{1/2}$，$f(\varepsilon_t)=|\varepsilon_t - b| - c(\varepsilon_t - b)$是 Pagan 和 Schwert（1990）引入的消息效应曲线。当$\lambda > 1$，Box-Cox 变换是凸的，当$\lambda < 1$，它是凹的。当$\lambda = v = 1$且$|c| \le 1$，（2.7）就成为特殊的 AGARCH 模型。当$\lambda = v = 1$且 b=c=0，就得到了 Taylor（1986）、Schwert（1989）的条件标准差模型。（2.7）中，当$\lambda = 0$、v=1 且 b=0 时，就产生了（2.4）中 p=q=1 时的指数 GARCH 模型。当$\lambda = v = 1$、b=0 且$|c| \le 1$，由（2.7）可得到标准差的 TGARCH 模型。如果$\lambda = v = 2$且 b=c=0，就产生（2.2）的 GARCH 模型。Engle 和 Ng（1993）提出的非线性非对称 GARCH 模型与$\lambda = v = 2$且 c=0 相对应，而当$\lambda = v = 2$且 b=0 的时候，就得到 Glosten-Jagannathan-Runkle（1993）提出的的 GARCH 模型。Higgings 和 Bera (1992)的非线性 ARCH 模型允许λ取任意值，且$v = \lambda$，b=c=0。Ding、Granger 和 Engle（1993）的非对称幂 ARCH（APARCH）模型允许λ取任意值且$v = \lambda$，b=0，$|c| \le 1$。Sentana （1991）的 QGARCH 模型未嵌入（2.7）的设定中。Hentschel（1994）指出，将现有的 GARCH 模型嵌入如（2.7）的一般设定强调了这些模型之间的关系，并提供了检验条件二阶矩函数形式嵌套假设序列的机会。Grouhy 和 Rockinger（1994）提出了通常称为滞后作用（hysteresis）的 GARCH（HGARCH）模型。在这个模型中，除了阈 GARCH 部分之外，还包括了收益对波动率的短期（几天）冲击和长期（几周）冲击。

Engle、Lilien 和 Robins（1987）提出了依均值的 ARCH（ARCH in Mean，ARCH-M）模型。在这个模型中，条件均值是以下过程的条件方差的函数：

$$y_t = g(z_{t-1}, h_t) + h_t^{1/2}\varepsilon_t \quad , \qquad (2.8)$$

其中z_{t-1}是前定变量的向量，g 是z_{t-1}的函数，而h_t由 ARCH（q）过程生成。当然，当h_t服从 GARCH 过程时，表达式（2.8）就是一个依均值的 GARCH 方程。最简单的 ARCH-M 模型为$g(z_{t-1}, h_t) = \delta h_t$。依均值的 GARCH 模型自然地产生于金融理论，如$g(z_{t-1}, h_t)$可表示某些资产的预期收益率，其中h_t为风险的测度。因此依均值的方程（2.8）能够反映风险和期望收益之间的权衡关系。Pagan 和 Ullah(1988)称这些模型为考虑风险的模型。

2.3 可选择的条件波动率模型

不以 ARCH 设定为基础的度量波动率的方法也已见诸文献。例如 French 等 (1987)根据股票日收益率平方值的平均数构造股票月收益率的方差，并对这些月方差拟合 ARMA 模型。利用高频数据估计低频观测值条件方差的过程不能有效地使用全部数据，同时，从第二步估计中得到的一般标准差可能是不合适的。然而，对于初步数据分析，这一过程计算的简单性和由 Schwert（1989）提出来的相关过程（即条件标准差是从条件均值的第一步估计开始，由残差度量），使其成为除复杂 ARCH 之外可选择的模型。

波动率的相关估计量可以从期间内的最高价和最低价得到。Parkinson（1980）证明，对于具有不变方差和连续时间参数的随机游走过程，方差的高－－低估计量比基于同样数量的期末（end-of-interval）观测值的样本方差更为有效。沿着这个思路，可以用波动率与买入－卖出价差之间的关系来构建收益的方差估计量[例如，参见 Bollerslev 和 Domowitz (1993)]。同样地，具有随机波动率的期权定价公式 [例如，参见 Melino 和 Turnbull (1990)] 的近期发展已经证实了期权价值和标的证券方差之间的正相关关系，这可以用来评定证券价格的波动率。最后，在给定时点

上资产之间价格收益信息的分布也可以用来量化市场波动率。

为设定条件方差的形式，就必须定义条件信息集，并选择一种函数形式来反映条件集合与条件方差之间的映射关系。通常情况下，条件信息集仅限于包括序列自身的过去值。简单的条件残差方差的两步骤估计量可以通过残差平方对其自身滞后值的回归得到[(参见 Davidian 和 Caroll（1987）)]。Pagan 和 Schwert（1990）证明 OLS 估计量是一致的，但不是有效的。这一两步骤估计量的作用是基准的计算，比较直接。跳跃模型或混合模型能够与条件方差的 GARCH 表达式结合，已经被用于描述波动率的时间变差、金融序列的肥尾和不对称性。在泊松跳跃模型中，假设出现异常信息时，收益就会出现跳跃。在 t 时出现的跳跃次数 n_t，产生于参数为 λ 的泊松分布。在跳跃次数为 n_t 的条件下，收益通常呈均值为 $n_t\theta$、方差 $\sigma^2=\sigma_\xi^2+n_t\sigma_\psi^2$ 的正态分布。参数 θ 代表跳跃规模（Size）的期望值，收益的条件均值和方差取决于 t 时的跳跃次数。假设 σ_ξ^2 的产生服从 GARCH 过程，就可以引入另外的时间依赖性。

在金融文献中，随机跳跃通常利用泊松过程的平均数来建模[参见 e.g.Ball 和 Torous （1985）、Jorion（1988）、Hsieh（1989）、Nieuwland 等以及(1991)Ball 和 Roma(1993)]。Vlaar 和 Palm（1993）将欧洲货币体系（EMS）周汇率的泊松跳跃过程与贝努里跳跃模型作比较，发现在大多数情形下，两个模型的表现非常相似。贝努里过程有一个优点，那就是当截去(cutting off)泊松跳跃过程的无穷和时，可以避免产生截断误差。

混合参数 λ 可以随时间变化。例如，Vlaar 和 Palm（1994）假设描述欧洲货币市场风险溢价的贝努里跳跃模型中，混合参数 λ 取决于相对于德国的通货膨胀的差异。

考虑时间依赖性的另一个方法是假设 t 时经济处于状态 1 的概率是不同的，它取决于经济在 t-1 时是处于状态 1 还是状态 2。Hamilton（1989）已经提出了这样的模型，并应用于研究汇率[(参见 Engle 和 Hamilton(1990)]、利率[(参见 Hamilton(1988)]和股票收益率[(参见 Pagan 和 Schwert(1990)]。

在 Hamilton 的基本模型中，有一个不可观测的状态变量 z_t，可取值 0 或 1。从 t-1 时的状态 j 转移到 t 时的状态 i 的概率 P_{ij} 是常数，给定 $P_{11}=p, P_{10}=1-p, P_{00}=q, P_{01}=1-q$。Pagan（1995）证明，$z_t$ 可演化为一个 AR（1）过程。在 Hamilton 的模型中，假设收益观测值 y_t 由下式生成：

$$y_t=\beta_0+\beta_1 z_t+(\sigma^2+\phi z_t)^{1/2}\varepsilon_t \quad , \tag{2.9}$$

其中，$\varepsilon_t \sim NID(0,\sigma^2)$。$y_t$ 在这两个状态中的期望值分别为 β_0 和 $\beta_0+\beta_1$。方差分别为 σ^2 和 $\sigma^2+\phi$。因而，模型中就生成高波动率和低波动率的状态。期望收益也会随不同的状态类型而改变。以 t-1 时状态为条件的收益条件方差可以表示为：

$$Var(y_t \mid z_{t-1})=[\sigma^2+(1-q)\phi](1-z_{t-1})+[p\phi+\sigma^2]z_{t-1} \quad 。 \tag{2.10}$$

很显然，(2.10）的条件方差表现出时间依赖性。

Hamilton 和 Susme（1994）通过假设扰动项为 ARCH 过程，推广了马尔可夫状态转换模型，称之为状态转换的 ARCH 模型（SWARCH）。在方程（2.9）中，SWARCH 模型的条件均值线性依赖于状态变量 z_t。

假设 y_t 的扰动项服从误差 $u_t = \sqrt{g_{st}}\tilde{u}_t$ 的 p 阶自回归过程，其中 $\tilde{u}_t$ 服从 Glosten 等（1993）的模型中有杠杆效应的 ARCH（q）过程，g_{st} 是随状态而变化的常数因子。假定新生 $\tilde{u}_t$ 服从均值为 0 的条件学生 t 分布。状态的转换由一个不可观测的马尔可夫链来决定。作者使用的数据是 1962 年 7 月 3 日到 1987 年 12 月 29 日在纽约股票交易所交易的价值加权的股票投资组合的周收益率，用不同的 ARCH 模型与 SWARCH 模型做了多达四种状态的比较。SWARCH 的设定有杠杆效应、服从低自由度的条件学生 t 分布，发现允许有四种状态的模型表现最好。Cai（1994）用类似的思路，利用两状态的 SWARCH 模型考察 1964 年 8 月到 1991 年 11 月的三个月期国库券月收益率的波动持续性问题。在过去的研究中，ARCH 过程的持续性可以解释为过程的条件方差的截距离散移位（discrete shift）的结果。状态更换发生的两个时期是 1974 年 2 月—1974 年 8 月的石油危机，以及 1779 年 9 月—1982 年 8 月的联邦储备银行政策的变化。

使用非参数方法可以得到不依赖于具体函数形式假定的条件方差的估计。Pagan 和 Schwert（1990）以及 Pagan 和 Hong（1991）使用了非参数核估计量和非参数灵活傅立叶形式（Flexible Fourier Form, FFF）估计量。

y_t 的条件矩核估计量用 $g(y_t)$ 来表示，具有有限数目的条件变量 x_t，如下所示：

$$\hat{E}[g(y_t)|x_t] = \sum_{s=1}^{T} g(y_s)K(x_t - x_s) / \sum_{s=1}^{T} K(x_t - x_s)\ , \qquad (2.11)$$

其中，K 是平滑数据的核函数。可以使用的核多种多样，普遍使用的是 Pagan 和 Schwert（1990）的正态核：

$$K(x_t - x_s) = (2\pi)^{-1/2} \,|H|^{-1/2} \exp[-\frac{1}{2}(x_t - x_s)' H (x_t - x_s)] \qquad (2.12)$$

H 是对角阵，它的第 k 个对角线元素等于带宽(band width) $\hat{\sigma}_k T^{-1/(4+q)}$，$\hat{\sigma}_k$ 是 $x_{kt}, k = 1, \cdots, q$ 的标准差，q 是条件集合的维数。

另一个可供选择的非参数估计量，是利用序列展开式得到条件方差的总体近似值。在现有的许多序列展开式中，Gallant（1981）提出的灵活傅立叶形式（FFF）在金融学中得到广泛的应用。条件方差用 $\hat{e}_t$ 的过去值（y_t 的回归残差）构造的低阶多项式或三角式的和来表达。这样，σ_t 设定为：

$$\sigma_t^2 = \sigma^2 + \sum_{j=1}^{L} \left\{ \alpha_j \hat{e}_{t-j} + \beta_j \hat{e}_{t-j}^2 + \sum_{k=1}^{2} \left[\phi_{jk} \cos\left(k\hat{e}_{t-j}\right) + \delta_{jk} \sin\left(k\hat{e}_{t-j}\right) \right] \right\} \qquad (2.13)$$

理论上，三角式的项应趋于无限，但实际上，依据显著性，通常不值得取高于两阶的形式。(2.13) 有个缺点，那就是 σ_t^2 的估计值可能为负数。Pagan 和 Schwert（1990）把（2.13）的估计量用于股票收益率的研究，其中 L=1。σ_t^2 的估计值大致不变，在 $\hat{e}_{t-1}$ 的大部分范围内与核估计、GARCH（1，2）及 FFF 估计的结果相似。只有当 $\hat{e}_{t-1}$ 是很大的正数或负数时*，估计量才会表现出不同的

* 指绝对值很大的负数——译者注

行为。对于负的$\hat{e}_{t-1}$，波动率的估计值急剧增大。同样，当用 F 联合检验时，（2.13）的三角项表现出高度显著性。

在股价下跌的时期，用核函数或傅立叶序列得到的条件波动率的非参数估计值与通过 GARCH、EGARCH 及 Hamilton 模型得到的参数估计值是不同的。特别地，超乎预期的大的负收益，将导致波动率的大幅度增加。对于大的冲击，参数估计调整得慢，而且这种冲击效应具有持续性。当非参数方法对大的反向冲击作出了高度的非线性反应时，参数方法则利用了它持续性的一面。虽然条件波动率的非参数估计量具有比 GARCH、EGARCH 及 Hamilton 模型的参数估计量高得多的解释能力，尤其是在解释非对称性方面，但与参数方法相比，它们是无效估计量。这表明把这两种方法结合起来，以获得比目前的的应用更丰富的模型设定，无效性可以得到改善。

其他的非参数方法也已经见诸文献。Gourieroux 和 Monfort（1992）提出如下形式的阶梯函数(step function)近似得到y_t和σ_t之间的未知关系：

$$y_t = \sum_{j=1}^{J}\alpha_j I_{A_j}(y_{t-1}) + \sum_{j=1}^{J}\beta_j I_{A_j}(y_{t-1})\varepsilon_t \quad , \tag{2.14}$$

其中，$A_j, (j=1,2,\cdots,J)$是y_{t-1}数值集合的一个分块，$I_{A_j}(y_{t-1})$是一个指示（indicator）变量，当y_{t-1}属于A_j时取值为 1，否则为 0。ε_t为白噪声。这个模型叫作定性阈自回归条件异方差模型（Qualitative Threshold ，QTARCH）。

如果状态 j 应用于变量y_{t-1}，则α_j和β_j分别为y_t的条件均值和方差。y_t的过程由马尔可夫链产生的定性状态变量$z_t = \left(I_{A_1}(y_t),\cdots,I_{A_j}(y_t)\right)$决定。例如，分块子集$A_1,\cdots,A_j$可能与金融市场扩张和收缩的不同阶段相对应。通过充分地细分$A_1,\cdots,A_j$，可以利用式（2.14）逼近y_t的条件均值和方差的更复杂的设定。通过加入 GARCH 项还可以选择更精粹的条件方差设定。α_j和β_j的准最大似然估计量是在状态 j 下计算出的样本均值和方差。QTARCH 模型用阶梯函数逼近条件均值和方差，Zakoian（1994）的 TARCH 模型则依赖于逐段线性逼近来估计条件方差函数。非参数核估计量平滑条件矩，而 FFF 估计量使用比分段线性函数或阶梯函数更平滑的函数来逼近条件矩。沿着类似思路，Engle 和 Ng（1993）用线性样条来估计对消息的反应形态。他们的方法称部分非参数法(partially nonparametric，PNP)，因为长期记忆成分作为参数建模而消息与波动率之间的关系却作非参数化处理。

在广泛应用于金融数据依存关系分析的半参数方法中，应当提及的是由 Gallant 和 Tauchen（1989）提出的、以具有高斯 VAR 首项的序列展开式为基础上的半参数（SNP）模型。

假设一个全部过去值已知的 N×1 维向量y_t的条件分布，只依赖于y_t有限的 L 个滞后值，用长度为 LN 的向量$x_{t-1} = \left(y'_{t-L}, y'_{t-L+1},\cdots,y'_{t-1}\right)'$来表示。这个过程包括近似$y_t$的条件密度，其中假定$x_{t-1}$由一个截断的 Hermite 展开式给出，该展开式是z_t的多项式与标准正态密度的乘积，z_t是y_t的中心化和尺度化值，$z_t = R^{-1}(y_t - b_0 - Bx_{t-1})$。截断的展开式是半参数模型。给定$x_{t-1}$条

件下z_t的条件 SNP 密度可以由下式近似得到：

$$f\left(z_t \mid x_{t-1}\right)=\frac{\left[\sum_{|\alpha|=0}^{k_t} a_\alpha\left(x_{t-1}\right) z_t^\alpha\right]^2 \varphi\left(z_t\right)}{\int\left[\sum_{|\alpha|=0}^{k_t} a_\alpha\left(x_{t-1}\right) u^\alpha\right]^2 \varphi(u) d u} \quad , \qquad (2.15)$$

其中φ代表标准高斯密度，$\alpha=\left(\alpha_1, \alpha_2, \cdots, \alpha_N\right)'$，$z^\alpha=\prod_{i=1}^N\left(z_i\right)^{\alpha_i}$，其度数（degree）是$|\alpha|=\sum_{i=1}^N\left|\alpha_i\right|$，$a_\alpha(x)=\sum_{|\beta|=0}^{K_x} \alpha_{\alpha \beta} x^\beta$，$\beta=\left(\beta_1, \beta_2, \cdots, \beta_{NL}\right)'$，$|\beta|=\sum_{i=1}^{NL}\left|\beta_i\right|$，$x^\beta=\prod_{i=1}^{NL}\left(x_i\right)^{\beta_i}$，$k_z$和$k_x$是正整数。给定$x_{t-1}$条件下的$y_t$的条件密度为$h\left(y_t \mid x_{t-1}\right)=f\left[R^{-1}\left(y_t-b_0-B x_{t-1}\right) \mid x_{t-1}\right] / \det(R)$。

正如 Gallant 和 Tauchen（1989）指出的那样，同时增加k_z和k_x，SNP 模型就会无限精确地近似于包括肥尾分布 （象 t 分布）、偏斜分布的一类模型。因为在闭合形式下，ARCH 模型的平稳分布是未知的，所以不能说 ARCH 模型属于上述类型。但是 ARCH 模型的平稳分布是肥尾的，仅有有限矩，类似 t 分布，在一定条件下，ARCH 模型和 SNP 模型的方差是有限滞后值的多项式。人们可能因而认为，当k_z和k_x很大时，ARCH 模型的的条件密度将无限接近于 SNP 模型。当 L 很大时，对于条件方差是无限滞后多项式的 GARCH 模型，这可能也是正确的。

对 ARCH 框架的另一个应用是假设方差的变化服从某种潜在过程(laternt process)。这导致了随机方差或称随机波动率（SV）模型的产生[参见 Ghijsels 等（1995）]。为了说明的方便，假设漂移(drift)参数是 0。Taylor（1986）提出了收益率y_t的简化 SV 模型：

$$y_t=\varepsilon_t \exp\left(\alpha_t / 2\right), \varepsilon_t \sim NID(0,1), \qquad (2.16)$$

$$\alpha_{t+1}=\alpha_0+\phi \alpha_t+\eta_t, \eta_t \sim NID\left(0, \sigma_\eta^2\right),$$

其中，随机变量ε_t和η_t是相互独立的。

Hull 和 White（1987）已使用过这个模型，例如，将它运用于外币期权的定价。Taylor（1986 和 1994）探讨了其时间序列特征。Taylor（1994）证明了 SV 模型的统计特性，并把这些模型表示为自回归方差（ARV）模型。估计 SV 模型出现的主要困难是该模型非线性，也非条件高斯(not conditionally Gaussian)。许多估计方法，象广义矩法（GMM）或准最大似然方法（QML）用于估计 SV 模型都无效。但是，依赖于以模拟技术为基础的方法能够进行贝叶斯估计或经典似然分析[参见 Kim 和 Shephard(1994)]。在近期，只有很少的研究比较了 GARCH 和 SV 方法在波动率建模中的表现。Ruiz（1993）在研究 1981 年 10 月 1 日到 1985 年 6 月 28 日期间英镑、德国马克、日元、瑞士法郎对美元的日汇率时，运用并比较了 GARCH（1，1）模型、EGARCH（1，0）模型和 ARV（1）模型。这三个模型的样本期内的表现非常相似。当用模型用于预测样本之外的波动率时，ARCH 模型表现出严重的偏倚而 SV 模型却没有出现这一情况。

Kim 和 Shephard（1994）研究表明，对于 1962 年 7 月 3 日到 1987 年 12 月 31 日以及 1962 年 7 月 11 日到 1992 年 12 月 30 日期间的 S&P500 指数的日收益率和周收益率，简单的一阶 SV

模型对数据的拟合程度与流行的 ARCH 模型一样好。Danielsson（1994）发现，对于 1980 年到 1987 年的 S&P500 指数的日数据，EGARCH（2，1）模型比 ARCH（5）、GARCH（1，2）、IGARCH（1，1，0）模型表现更好。同时，它还优于用模拟最大似然法估计的简单 SV 模型。动态 SV 模型与 EGARCH 对数似然值之间的差额是 25.5，支持了 4 个参数的 SV 模型优于 5 个参数的 EGARCH 模型的结论。

2.4 多变量 GARCH 模型

除了 SNP 模型之外，2.2 节和 2.3 节中介绍的模型都是单变量的。在资产定价和投资组合配置中，许多问题的分析都需要用到多变量结构。

考虑一个 N×1 向量的随机过程{ y_t }，并把它写成：

$$y_t = \Omega_t^{1/2} \varepsilon_t \quad , \tag{2.17}$$

其中，ε_t 是一个 N×1 的 i.i.d.向量，$E\varepsilon_t = 0$，$\mathrm{Var}(\varepsilon_t) = \mathrm{I}_\mathrm{N}$，$\Omega_t$ 是 t 时可得信息条件下 y_t 的 N×N 协方差矩阵。

Bollerslev、Engle 和 Wooldridge（1988）假设，在一个多变量线性 GARCH（p,q）模型中，Ω_t 由误差交叉平方项的滞后值和 Ω_t 滞后值混合的线性函数给出：

$$vech(\Omega_t) = \alpha_0 + z\sum_{i=1}^{q} A_i vech(\varepsilon_{t-i}\varepsilon'_{t-i}) + \sum_{i=1}^{p} B_i vech(\Omega_{t-i}) \quad , \tag{2.18}$$

其中 vech(.)表示一个把 N×N 矩阵的下半部分排列成一个 N（N+1）/2×1 向量的算子。在（2.18）中，α_0 是 $N(N+1)/2$ 的向量，A_i 和 B_i 是 $N(N+1)/2$ 阶矩阵。未知参数的个数等于 N(N+1)[1+N(N+1)(p+q)/2]/2。在实践中，需要使用一些简化的假设以使之更简约。例如，Bollerslev 等（1988）在应用对角 GARCH（p,q）模型时，假设矩阵 A_i 和 B_i 是对角阵。其他一些代表性的研究结果，如 Baillie 和 Bollerslev（1990）以及 Vlaar 和 Palm（1993）使用的常数条件相关模型，都假设条件方差是 GARCH 过程。

参数化（2.18）的条件是确保对于 ε_t 的任何值，Ω_t 都为正定矩阵，但这一点在实际中很难验证。Engle 和 Kroner（1995）提出了参数化多变量 GARCH 过程的方法，将它称为 BEKK（Baba、Engle、Kraft 和 Kroner）表达式：

$$\Omega_t = C_0^{*'} C_0^* + \sum_{k=1}^{K}\sum_{i=1}^{q} A_{ik}^{*'} \varepsilon_{t-i}\varepsilon'_{t-i} A_{ik}^* + \sum_{k=1}^{K}\sum_{i=1}^{p} G_{ik}^{*'} \Omega_{t-i} G_{ik}^* \quad , \tag{2.19}$$

其中 C_0^*，A_{ik}^* 和 G_{ik}^* 是 N×N 阶参数矩阵，C_0^* 是三角阵，求和界限 K 决定了过程的通用性。(2.19）中的协方差矩阵在弱条件(weak conditions)下为正定矩阵。这个表达式也具有充分的一般性，因为它包括了(2.18)中所有正定对角表达式和大部分正的 vec 表达式。就参数个数而论，(2.19)比(2.18）更简约。假定在相当一般的情况下，发现两种参数化的方法等价，BEKK 参数化的方法会更受欢迎，因为此时正定性易于保证。

Engle、Ng 和 Rothschild（1990）提出用因子—ARCH 模型作为作为资产超额收益率条件协方差矩阵更简约的结构。这些模型体现了这样一个概念，即金融资产的风险可以分解为有限数目的共同因子 f_t 和一个特定资产的扰动项。套利定价理论（APT）产生了单因子结构，虽然 APT 并未意味着因子数目是有限的。Engle、Ng 和 Rothschild（1990），将因子—ARCH 模型应用于利率风险的建模，Ng 等（1992）在相关的论文中研究了美国股票市场的资本资产定价模型（CAPM）的风险溢价和异常现象。Diebold 和 Nerlove（1989）将单因子模型应用于汇率中， King、Sentana 和 Waadhwani（1994）用单因子模型分析了各股票市场之间的联系。

因子模型如下所示：

$$y_t = \mu_t + Bf_t + \varepsilon_t \quad , \tag{2.20}$$

其中，y_t 是 N×1 的收益向量，μ_t 是 N×1 阶期望收益向量，B 是 N×k 的因子载荷矩阵，f_t 是条件协方差矩阵为 Λ_t 的 k×1 的因子向量，ε_t 代表条件协方差为 Ψ_t 的 N×1 的特质（idiosyncratic）冲击因子向量。各因子和特质冲击是不相关的。y_t 的条件协方差矩阵由下式给出：

$$\Omega_t = B\Lambda_t B' + \Psi_t \tag{2.21}$$

当 Ψ_t 是常数且 Λ_t 具有常数（可能为 0）的非对角元素时，协方差矩阵 Ω_t 可以由下式给出：

$$\Omega_t = \sum_{i=1}^{k} b_i b_i' \lambda_{it} + \Psi \quad , \tag{2.22}$$

其中 b_i 代表 B 的第 i 列，Ψ 集中了 Λ_t 的非对角线元素，以及 ε_t 的协方差矩阵中的常数元素。正如 Engle 等（1990）所指出的，根据观测，（2.22）的模型等价于一个有着常数 λ 而 b 随时间变化的类似模型。因子模型（2.22）还隐含着，如果 k<N，可以构造 N－k 种资产投资组合，即 y_t 的线性组合，它有常数方差。而有 k 种投资组合是以 λ_{it} 加一个常数为条件方差。

必须详加说明条件方差的生成过程，（2.20）的因子模型才得以完成。例如，可以假设 λ_{it} 由单变量 GARCH 过程生成。Diebold 和 Nerlove（1989）运用单因子模型研究 1973 年 7 月至 1985 年 8 月七种货币对美元汇率周数据的对数差分，他们假设单一的共同因子的方差为 $\lambda_t = \alpha_0 + \theta \sum_{i=1}^{12} (13-i) f_{t-i}^2$。应注意的是，它们的协方差矩阵是 7×7 的，但只包括 9 个未知参数， 如 Ψ , α_0 和 θ。通过将线性衰减模式加之于 ARCH 系数，大大减少了需要估计的参数个数。用 GARCH（1，1）设定替代，可产生几何衰减的 ARCH 系数。

Engle 等（1990）提出了另一可选择的方法。他假设每一有 k 个有代表因子的投资组合的收益服从 GARCH 过程。当 i=1,⋯,k 时，第 i 个投资组合的条件方差由下式给出：

$$\phi_i' \Omega_t \phi_i = \varpi_i + \alpha_i (\phi_i' \varepsilon_{t-1})^2 + \beta_i \phi_i' \Omega_{t-1} \phi_i \quad , \tag{2.23}$$

为了简便，假定是 GARCH（1，1）模型，ϕ_i 是 N×1 投资组合权重向量。投资组合的条件方差与 λ_{it} 的条件方差的不同之处仅是常数项，即 $\phi_i'\Omega_t\phi_i = \lambda_{it} + \phi_i'\Psi\phi_i$。把它和（2.23）一起代入（2.22），以便表达（2.22）中根据条件投资组合方差得到的条件协方差矩阵 Ω_t。应注意的是，$\varepsilon_i'b_i = 1$ 且 $\phi_i'b_j = 0, j \neq i$。虽然因子—GARCH 模型在理论上显示出有吸引力的特征，但它的估计需用到高度非线性方法。在这些方法中，Lin（1992）考虑了最大似然估计法。模型在能够估计之前因子组合未直接观测到，所以还必须解决识别问题[参见 Sentanta（1992）]。特别是需要识别代表组合的因子。在一些例子中，有必要假设代表组合的因子已知或可观测。例如 Engle 等（1990）解释了 1964 年 8 月到 1985 年 11 月期间，到期日从 1 个月用到 12 个月的政府国库券月收益率以及 NYSE-AMSE 股票的加权指数。他们选择了两个组合代表因子，其中一个对每种国库券有相同权数而股指权数为 0；另一个的国库券权数为 0，而所有的权数都在股指上。可观测的因子代表的组合的模型可用两阶段方法得到一致估计。首先估计组合的单变量模型，使用第一阶段的估计结果可以得到因子载荷的一致估计，只要单个资产的方差与因子方差存在线性关系，系数等于因子载荷的平方。

King 等（1994）利用 1970 年 1 月到 1988 年 10 月期间的 16 个股票市场的美元月超额收益率数据，用最大似然法估计（2.20）形式的多变量因子模型。他们假设风险溢价 μ_t 可表示为 $\mu_t = B\Lambda_t\tau$，其中 Λ_t 是一个对角阵，τ 是一个 k×1 向量，代表每个因子风险价格的常参数。King 等（1994）考虑了 k=6，其中 4 个可观测因子，2 个不可观测因子的模型。可观测的因子代表对资产收益不可预料的冲击。这些冲击是作为从一个四因子模型提取的共同因子来估计的，将四因子模型应用于 x_t——10 个可观测的宏观经济变量的向量自回归——的残差，得到共同因子的估计。假定共同项和特质项的方差服从单变量 GARCH（1，1）过程，其中，因子过去值的平方由它们在给定可得信息集下的线性投影来替代。应注意的是，当因子 GARCH 模型的协方差矩阵依赖于先验未观测值时，收益成分具有条件随机波动率的代表性[参见 Anderson（1992）及 Harvey 等（1992）]。

一个主要的发现是，联邦股票市场与其时间变差之间的协方差，只有一小部分能够用可观测的因子来解释。条件二阶矩在很大程度上是由不可观测因子来解释的。这一发现强调了不可观测因子模型在解释市场内部波动率以及市场之间的波动率溢出时是很有用的。King 等（1994）也阐述了因子模型在解释 16 维多变量时间序列二阶矩的时间依赖关系中的适用性和可行性。虽然联合估计大约 200 个参数的因子模型是可能的，作者却不得不在第一阶段单独估计 x_t 的向量自回归。由于多变量因子—GARCH 模型参数空间的维数通常较高，相对于以似然原则为基础的完全联合估计方法，两阶段估计方法是可行的选择。

2.5 条件方差的一致性

对于高频时间序列数据，用（2.2）的 GARCH（p,q）过程估计的条件方差经常表现出持续性，也就是 $\sum_{i=1}^{p}\beta_i + \sum_{i=1}^{q}\alpha_i$ 接近于 1。当总和等于 1，就产生了 IGARCH 模型。这意味着当预测任何时间跨度(all horizons)的条件方差时，当前的信息仍然是重要的。非条件方差在这个情形中不存

在。Bollerslev（1986）已经证明，在正态情况下，GARCH 过程（2.2）是广义平稳的。当且仅当 $\sum_{i=1}^{p}\beta_i+\sum_{i=1}^{q}\alpha_i<1$ 时，非条件方差 $\mathrm{var}(y_t)=\alpha_0(1-\sum_{i=1}^{p}\beta_i-\sum_{i=1}^{q}\alpha_i)^{-1}$，$\mathrm{cov}(y_t,y_s)=0, t\neq s$。Nelson（1990a）、Bougerol 和 Picard（1992）证明了 IGARCH 模型是严平稳、遍历但非协方差平稳的。

类似地，正如 Bollerslev 和 Engle（1993）所证明的，当且仅当特征多项式 $\det[I-A(\lambda^{-1})-B(\lambda^{-1})]=0$ 的根在单位圆内时，多变量 GARCH（p,q）过程（2.18）是协方差平稳的。在这一情形中，方差将不具有持续性。另一方面，如果一些特征值落在单位圆上，条件协方差矩阵所受的冲击对所有时间跨度的预测仍然很重要。如果特征值在单位圆之外，对协方差矩阵的冲击效应就会随时间急剧变化（explode）。应注意的是，如 Engle 和 Kroner（1995）指出的那样，上述特征多项式根的条件同样可应用于BEKK模型(2.19)。在许多运用单变量GARCH(p,q)模型对金融数据进行的实证研究中，发现估计参数的和接近于 1。详细的文献综述可参见 Bollerslev、Chou 和 Kroner（1992）。如果投资组合和 ε_t 是协方差平稳，则具有（2.23）形式的 GARCH（p,q）过程的 k 因子多变量模型（2.20），其因子也是协方差平稳的。

和变量间的协整概念一致，Bollerslev 和 Engle（1993）提出了方差共同持续性（co-persistence in variance）的定义。其基本思想是，虽然许多时间序列的方差可能表现出持续性。但同时，一些变量线性组合的方差可能表现出不具有持续性。Bollerslev 和 Engle（1993）推导出多变量 GARCH（p,q）过程的方差具有共同特续性的充要条件。在实际中，方差共同持续性允许人们从非平稳收益波动率的资产中构造出平稳波动率的投资组合。

多变量 GARCH 模型单位根的发现推动了因子—ARCH 模型的新进展。Engle 和 Lee（1993）构造了 King 等（1994）形式的因子模型，模型中考虑了波动率的永久成分 IGARCH（1，0，1）和暂时成分 GARCH（1，1）。

Engle 和 Lee（1993）将成分模型的几个变形应用于 CRSP 价值加权指数和 14 种美国大公司股票的日收益率，样本期为 1962 年 7 月 1 日到 1991 年 12 月 13 日。主要实证结果是：个股收益波动的持续性归因于市场波动率（假设它是一个共同因子）和个股特质波动持续性。这一结果意味着，当假定市场冲击不影响特质波动率时，股票收益的波动率与市场波动率具有共同特续性的假设要被拒绝。

Palm 和 Urbain（1995）利用具有可观测因子的因子——成分——GARCH 模型 1982 年 2 月到 1995 年 8 月期间欧洲、远东和北美股价指数收益率的日观测值，发现共同因子和特定因子的波动率具有显著的持续性。

虽然对因子——成分——GARCH 模型的应用尚处在初始阶段，有关收益波动的持续性[例如，参见 French、Schwert 和 Stambaugh（1987）、Chou（1988）、Pagan 和 Schwert（1990）、Ding 等(1993)以及 Engle 和 Gonzalez-Rivera(1991)]、共同因子和/或特质因子的实证研究结果提出了许多重要问题。例如：有关波动持续性的发现是否符合文献中经常用到的资产收益的平稳性假设？金融理论是否预示着非平稳波动将导致非平稳的资产收益？在波动率和收益序列中存在着的持续性的确切形式是什么？它是否应当作为条件方差永久成分的一个单位根来建模？或者允许它为分数单整？或者应当象诸如 Cai(1994)Hamilton 和 Susmel（1994）那样，按状态转换来建模？越来越多的证据表明收益序列表现出分数单整[例如，参见 Baillie（1994）]。由于现有多数检验方法的低效力，要完全根据观察来区分单位根引起的持续性和分数单整引起的持续性是十分困难的。

3. 统计推论

3.1 估计与检验

通常用最大似然法（ML）或准最大似然法（QML）来估计 GARCH 模型。在一些实际应用中，也用到广义矩法（GMM）[例如，参见 Glosten 等(1993)]。随机波动率模型一般用 GMM 方法估计。更近期的间接推断法[例如，参见 Gourieroux 和 Monfort(1993)以及 Gallant 等(1994)]也已经被提倡并应用于估计随机波动率模型。贝叶斯方法已发展用于波动率模型[例如，参见 Jacquier 等（1994）的随机波动率模型估计以及 Geweke（1994）对随机波动率 GARCH 模型的估计]。为了简便的原因，我们讨论ε_t服从 IN（0，1）的假定下，（2.1）和（2.2）中 GARCH（1，1）模型的 ML 估计。记y_t的 T 个观测值为$y=(y_1,y_2\cdots,y_T)'$，T 个观测值的对数似然函数 L 可写为：

$$L(y\,|\,\theta)=\sum_{t=1}^{T}L_t \quad , \tag{3.1}$$

其中$L_t=c-\dfrac{1}{2}\ln h_t-\dfrac{1}{2}y_t^2/h_t$，$\theta=(\alpha_0,\alpha_1,\beta_1)'$，$h_1=\sigma^2=\alpha_0/(1-\alpha_1-\beta_1)$，$h_t$由（2.2）给出，t>1。

给定参数向量θ的初始值，通过递归计算h_t，t=1,2⋯,T，并将它的值代入（3.1）可求出对数似然函数的值。标准数值算法可用于计算（3.1）的最大值。众所周知，在一定的正则性条件下，例如 Crowder（1976）条件，使 L 最大的θ的估计量$\hat{\theta}_{ML}$是一致、渐进正态和有效的：

$$\sqrt{T}(\hat{\theta}_{ML}-\theta)\overset{a}{\sim}N(0,Var(\hat{\theta}_{ML})) \quad , \tag{3.2}$$

其中，$Var(\hat{\theta}_{ML})=-[T^{-1}\sum_{t=1}^{T}E\partial^2L_T/\partial\theta\partial\theta']^{-1}$。$\hat{\theta}_{ML}$的渐进协方差矩阵可以通过与（3.1）相关联的 Hessian 矩阵的逆一致地估计，在$\hat{\theta}_{ML}$处求值。Lumsdaine（1992）在$E[\ln(\alpha_1\varepsilon_t^2+\beta_1)]<0$的条件下证明了 GARCH（1，1）和 IGARCH（1，1）模型 ML 估计量的一致性和渐进正态性。不要求ε_t存在有限四阶矩。和条件均值中有单位根的模型不一样，不论条件方差有或没有单位根，模型的 ML 估计量具有相同的极限分布。

如同 Weiss（1986）证明了具有 ARCH 误差的时间序列模型，以及 Bollerslev 和 Wooldridge（1992）、gourieroux（1992）证明了 GARCH 过程，通过使正态分布的对数似然函数（3.1）最大化，可以得到θ的准 ML（quasi-ML）估计量或拟 ML（pseudo-ML）估计量，虽然真实的概率密度函数是非正态的。在正则性条件下，QML 估计量有以下的渐进分布：

$$\sqrt{T}(\hat{\theta}_{QML}-\theta)\sim N(0,B^{-1}AB^{-1}) \tag{3.3}$$

其中，$A=E_0[\partial L_t/\partial\theta\bullet\partial L_t/\partial\theta']$是 L 的得分向量的协方差矩阵。$B=-E_0[\partial^2L_t/\partial\theta\partial\theta']$，$E_0$代表在数据的真实概率密度条件下的期望。当然，如果后者是正态分布，（3.2）和（3.3）中的渐进分布就是相同的。Lee 和 Hansen（1994）证明了高斯 GARCH（1，1）模型的 QML 估计量的渐

进正态性。用条件标准差来衡量的扰动项，并不要求服从正态分布，也不要求是独立的。如果尺度化（scaled）扰动项的条件四阶矩是有界的，当$\alpha_1+\beta_1=1$时，GARCH 过程可能是单整的，$\alpha_1+\beta_1>1$，它甚至是爆发(explosive)的。Bollerslev 和 Wooldridge（1992）在模拟研究中发现，有限样本中，当对称地偏离条件正态性时 QML 估计量近似于精确的 ML 估计量。对于非对称的真实条件分布，与精确 ML 估计量相比，无论是大样本还是小样本，QML 有效性的损失都是相当大的。Engle 和 Gonzalez-Rivera（1991）提出的半参数密度估计使用先前平滑的线性样条，将成为除了 QML 方法之外的另一种可选择的具有吸引力的方法。

对估计 GARCH 模型的 ML 方法和 QML 方法可以作一些评论。首先， GARCH 在非条件分布中产生了肥尾，虽与条件正态性结合，但它并不能完全解释许多金融数据中表现出来的超峰度。许多作者还使用了自由度待估计的学生 t 分布。其他已经在 GARCH 模型估计中使用过的密度包括正态—泊松混合分布[例如，参见 Jorion（1988）和 Nieuwland 等(1991)]、正态—对数正态混合分布[例如，参见 Hsieh（1989）]、广义误差分布[例如，参见 Nelson(1991)]和贝努里—正态混合分布[例如，参见 Vlaar 和 Palm(1993)]。De Vries（1991）提出用一个条件平稳分布的类 GARCH 过程，这个过程用于对波动率群集建模，它具有肥尾和非条件平稳分布。

第二，对一些模型，例如正态的 ARCH 扰动模型，信息矩阵是分块对角阵[例如，参见 Engle(1982)]，由于回归系数和 ARCH 参数可以在不影响渐进有效性的条件下单独估计，所以其意义非常重大。而且，它们的方差也可以分别得到。Linton（1993）扩展了这一结果，他认为，当误差服从平稳的 ARCH（q）过程且未知的条件密度以零为对称时，条件均值的参数在 Bickel 意义上是合适的。换句话说，使用以正态密度函数为基础的核方法来估计未知的得分函数，产生条件均值的参数估计，这一条件均值与建立在真实分布基础上的 ML 估计量有相同的渐进分布。对 ARCH-M 模型，用对角化分块矩阵是不合适的，因为序列的条件均值依赖于条件方差过程的参数。同样，对 EGARCH 扰动过程，信息矩阵的分块对角化也是不合适的。

Gourieroux 和 Monfort（1993）提出了间接推断。当 QML 或 ML 的运用有困难、但还是可以从数据中估计某些关注参数的函数时，Gallant 等（1994）提出的有效矩方法显得很有吸引力。

Engle 和 Lee（1994）将间接估计方法应用于估计随机波动率扩散模型。作为出发点，他们估计了 1990 年 9 月到 1991 年 1 月 S&P500 指数日收益率的 GARCH（1，1）模型，得到 QML 估计量$\hat{\theta}$，并用于估计资产价格p_t和条件方差σ_t^2的现有扩散模型参数：

$$\begin{aligned}
&(a) \quad y_t=\mu d_t+\sigma_t dw_{yt} \\
&(b) \quad d\sigma_t{}^2=\phi(\varpi-\sigma_t^2)d_t+\xi\sigma_t^{\delta}dw_{\sigma t} \\
&(c) \quad correl(dw_y,dw_\sigma)=\rho
\end{aligned} \qquad (3.4)$$

其中，$y_t=dp_t/p_t$，dw_y和dw_σ服从维纳过程，使用的关系式与 GARCH 模型及扩散模型的条件一阶矩、条件二阶矩是相匹配的[参见 Nelson（1990b）]：$\varpi=\alpha_0,\phi=(1-\alpha_1-\beta_1)dt,\xi=\alpha_1\sqrt{(\kappa-1)}dt,\delta=1$，$\kappa$是 GARCH 模型冲击的条件峰度。

Gallant 等（1994）使用以两个辅助模型得分值为基础的间接方法来估计随机波动率模型。这两个辅助模型都假定其 SNP 密度由（2.15）给出。当 SNP 密度都具有条件同方差和非高斯新生的 ARCH 形式时，就把它叫做非参数 ARCH 模型，因为它与 Engle 和 Gonzalez-Rivera（1991）所研究的非参数 ARCH 过程十分相似。第二个模型放弃同方差性约束，称为完全非参数设定。SNP 模型是用 QML 估计的。

Gallant 等（1994）用 1928 年—1987 年 S&P 合成指数的日观测值估计单变量模型，用 1977 年—1992 年的日观测值估计 S&P NYSE 指数、德国马克对美元汇率和三个月期欧元利率的三变量模型。他发现，对于股价和利率，随机波动率模型可与非参数 ARCH 得分的 ARCH 部分相匹配。但是，它不能匹配新生分布中的矩，对于汇率，随机波动率模型并不能拟合其中的 ARCH 部分。

已经有许多文献对 ARCH（q）模型的存在性作了检验。简便而又广泛使用的是 Engle 提出的拉格朗日乘数（LM）检验，其零假设为 $H_0:\alpha_1=\alpha_2=\cdots=\alpha_q=0$，备择假设为 $H_0:\alpha_1\geq 0,\cdots,\alpha_q\geq 0$ 且至少有一个严格不等式成立。

$$LM=\frac{1}{2}f_0'z(z'z)^{-1}z'f_0 \quad , \tag{3.5}$$

其中 $z_t=(1,y_{t-1}^2,\cdots,y_{t-q}^2)'$，$z=(z_1,\cdots,z_T)'$ 且 f_0 是 (y^2/α_0-1) 的列向量。

$LM=TR^2$ 是一个渐进等价的统计量。这里，R^2 是 f_0 和 z 的复相关系数的平方。T 是样本容量。同时 R^2 也是 y_t^2 对截距项及 y_t^2 的 q 期滞后值回归的可决系数。正如 Engle（1982）所指出的，双边 LM 检验具有自由度为 q 的渐进 χ^2 分布。Demos 和 Sentana（1991）报告了 LM 单边检验的临界值，它对非正态是稳健的。构造 GARCH 分布的 LM 检验的一个困难，如 Bollerslev（1986）指出，在于需要用到逆矩阵的分块信息矩阵是奇异的。这是由于在零假设下，GARCH（1，1）模型中的 β_1 不可识别。Lee（1991）已提出如何避免这一困难，他还证明，ARCH 和 GARCH 误差的 LM 检验是相同的。

Lee 和 King（1993）推导出基于局部平均最大功效（LMMP）检验的得分（LBS）检验以验证 ARCH 和 GARCH 扰动的存在。这个检验是基于在零假设下的评估得分之和以及多余参数(nuisance parameters)用它们的 ML 估计量代替。当没有多余参数时，检验就是 LMMP。通过将得分和除以大样本标准误，得分和就标准化了。所得到的检验统计量具有渐进 N（0，1）分布。用于检验 ARCH（q）过程的统计量也可用来检验 GARCH（p,q）过程。在小样本中，LBS 检验比 LM 检验有更好的功效，至少已发现它的渐进临界值是同样精确的。

Wald 和似然比（LR）准则可用于如下的检验：零假设为条件同方差，即备择假设为 GARCH（1，1）。与 $H_0:\alpha_1=0$ 且 $\beta_1=0$，$H_1:\alpha_1\geq 0$ 或 $\beta_1\geq 0$（至少有一个严格不等式成立）相关联的检验统计量并不服从自由度为 2 的 χ^2 分布，因为在 H_0 条件下，实际参数值不依赖于参数空间边界的标准假定不成立。能够证明自由度为 2 的 χ^2 分布的 LR 检验是很保守(conservative)的[参见 Kodde 和 Palm(1986)]。同样，缺少上面提出的参数识别会导致 Wald 和 LR 的标准检验无法进行。这些 ARCH 统计量检验针对的是条件异方差的特定形式，然而，许多检验是设计来检验对独立、同分布随机变量的总偏离的。例如，Brock、Dechert 和 Scheinkman（1987）的 BDS 检验是用于检验一般的非线性相关。它拒绝 ARCH 备择假设的功效与 LM-ARCH 检验相似[参见 Brock、Hsieh 和 Lebaron（1991）]。对于其它备择假设，BDS 检验的功效可能更高。Bera 和 Lee（1993）把 white 信息矩阵（IM）准则用于自回归扰动项的线性回归模型，这导致了 Engle 的 LMARCH

检验的推广，其中 ARCH 被设定为随机系数自回归模型。许多作者已经提及可以给出 ARCH 的随机系数解释[参见 Tsay(1987)]。Bera、Lee 和 Higgings（1992）指出，如果每次仅涉及某一设定，而不是将它们结合起来考虑，那是很危险性的，他还给出了同时分析自相关和 ARCH 的框架。Diebold（1987），以一种令人信服的方式举例说明这一框架是需要的。他还证明，在 ARCH 过程中，序列相关的标准检验将导致过度拒绝零假设。要注意的是，ARCH 的存在可以用许多方法来解释，比如非正态（不对称 ARCH 的超峰度和偏斜）[例如，参见 Engle(1982)]和非线性[参见 Higgings 和 Bera(1992)]。

最近， Bollerslev 和 Wooldrige（1992）发展了稳健的 LM 来检验均值和方差联合参数化的充分性。他们的检验以对约束 QML 估计量求值的对数似然函数梯度为基础，可以用简单的辅助回归计算出来，只需要求条件均值与方差的的一阶导数。作者以模拟结果显示在大多数情况下，稳健检验统计量要比非稳健（标准）Wald 和 LR 检验更受欢迎。

这个结论与 Lumsdaine（1995）的发现一致。他在对有限样本的 ML 估计量性质和相关检验统计量的模拟研究中比较了 GARCH（1，1）和 IGARCH（1，1）模型。发现当估计的 t 统计量非常接近于渐进分布时，参数估计量对于有限样本容量是偏斜的，Wald 检验有最佳的基准（Size），标准 LM 检验基准太大，但对于可能的非正态的稳健检验统计量表现较好。

有关文献中提出了各种各样的模型诊断。例如 Li 和 Mak（1994）检验了高斯过程的标准残差平方自相关的渐进分布，并用 ML 方法估计了这一过程中随时间变化的条件均值与方差。然后用残差除以它们的条件标准差再减去它们的样本均值，就将残差标准化了。这一过程的条件均值与方差可以是 t 时可得信息的非线性函数。这些函数假定有连续二阶导数。当数据的生成过程是 ARCH(q)时，建立在 r 至 M 阶标准化残差平方自相关基础上的 Box-Pierce 混成检验(portmanteau test)服从自由度为 M－r（当 r>q）的 χ^2 分布。这些类型的诊断对于检查模型的适当性非常有用。

多变量 GARCH 模型可能会产生一些特殊形式的假设检验问题。例如，GARCH 可以是许多时间序列的共同特征。Engle 和 Kozicki（1993）定义了一组时间序列的共同特征，它们对于其非零线性组合不具有该特征的那些序列是共同特征。作为一个例子，考虑（2.20）的因子—ARCH 模型的双变量形式，它具有一个因子和一个常数特质因子的协方差阵。如果 f_t 的方差服从 GARCH 过程， y_{it} 序列也是 GARCH 过程，但线性组合 $y_{1t}-b_1/b_2y_{2t}$ 却有一个常数条件方差。在这个例子中，序列 y_{1t}、y_{2t} 享有一个共同特征，即随时间变化的条件方差的共同因子形式。Engle 和 Kozicki（1993）提出共同特征检验。Engle 和 Susmel（1993）应用这一程序检验了国际股票市场 ARCH 的共同特征。方法如下：首先，检验 ARCH 在单独时间序列中的存在性；第二，如果 ARCH 的效应在两个时间序列中都是显著的，考虑线性组合 $y_{1t}-\delta y_{2t}$，将其平方值对滞后平方值和序列 y_{it} 直至 q 期的滞后交叉乘积做回归，并对系数 δ 最小化 $TR^2(\delta)$。如果考虑的不是两个时间序列，而是一组的 k 个序列，δ 成为一个(k-1)×1 向量。Engle 和 Kozicki（1993）证明，关于 δ 的 $TR^2(\delta)$ 最小化的检验统计量具有 χ^2 分布，它的自由度等于回归中包含的滞后平方值的个数减去（k－1）。Engle 和 Susmel（1993）将检验应用于 1980 年 1 月到 1990 年 1 月世界上 18 个主要股票市场的股票市场指数的周收益率。他们发现有两组国家，一组是欧洲国家，另一组是远东国家表现出相似的随时间变化的波动率，这样，共同特征检验证实了每一组中共同因子—ARCH 结构的存在。

4. 统计特征

在这一节，我们将总结 GARCH 模型的主要统计特征，并给出合适的参考文献。

4. 1 矩

1986 年，Bollerslev 证明，在条件正态下，当且仅当$\alpha(1)+\beta(1)<1$时，（2.2）的 GARCH 过程是宽平稳的，并且$Ey_t=0$，$\mathrm{var}(y_t)=\alpha_0[1-\alpha(1)-\beta(1)]^{-1}$，当$t\neq s$时，$\mathrm{cov}(y_t,y_s)=0$。对于（2.2）中的 GARCH（1，1）模型，存在 2r 阶矩的充要条件是当$\alpha_0=1$且$\alpha_j=\prod_{i=1}^{j}(2i-1), j=1,2,\cdots$时，$\sum_{j=0}^{r}\binom{r}{j}\alpha_j\alpha_1^j\beta_1^{r-j}<1$。Bollerslev（1986）还提出了当 p=q=1 时，y_t的偶数矩的递归公式。如果条件正态 GARCH（1，1）变量存在四阶矩，那它就等于$Ey_t^4=3(Ey_t^2)^2[1-(\beta_1+\alpha_1)^2]/[1-(\beta_1+\alpha_1)^2-2\alpha_1^2]$。由于正态分布的对称性，如果存在奇数矩，则奇数矩等于 0。这些结论扩充了 Engle（1982）给出的 ARCH（q）过程的结果。上面给出的条件对严平稳是充分的，但不是必要的。

Krengel（1985）证明，向量 ARCH 过程y_t的严平稳等价于$\Omega_t=\Omega(y_{t-1},y_{t-2},\cdots)$是可测的，而且几乎必定（a.s.）$\Omega_t\Omega_t'<\infty$ [参见 Bollerslev 等(1994)]。对于某些 r>0，矩是有界的，即$E[trace(\Omega_t\Omega_t')^r]$，隐含着几乎必定（a.s.）$trace(\Omega_t\Omega_t'$）$<\infty$ 。Nelson（1990a）已证明，对于（2.2）的 GARCH（1，1）模型，当且仅当$E[\ln(\beta_1+\alpha_1\varepsilon_t^2)]<0$，$\varepsilon_t$是 i.i.d. (不必是条件正态)，$y_t$是严平稳的，$y_t^2$是非退化的。这个要求比$\alpha_1+\beta_1<1$要弱得多。他还证明没有漂移的 IGARCH（1，1）模型几乎必定（a.s.）收敛到 0，而当存在正的漂移时，它是严平稳和遍历的。一般单变量 GARCH（p,q）过程的扩展结果也已经由 Bougerol 和 Picard（1992）得到。

4. 2 GARCH 和连续时间模型

GARCH 模型是非线性随机差分方程，它比金融理论文献中对随时间变动的波动率建模时采用的随机微分方程更容易估计。在实际中，观测值通常是在离散点被及时记录下来的，所以统计推断必然要使用离散时间模型或近似于连续模型的离散时间模型。Nelson(1990b)推导出，当观测值之间的间隔长度 h 趋于 0 时，ARCH 过程从随机差分方程收敛到随机微分方程应具备的条件。他将结果应用于 GARCH（1，1）和 EGARCH 模型。Nelson（1992）研究了错误设定 ARCH 模型时产生的条件协方差矩阵估计的性质。以离散时点间隔 h 对扩散过程进行观测，发现随着$h\downarrow 0$，基于 GARCH（1，1）或 EGARCH 模型的条件瞬时（conditional instantaneous）协方差矩阵的估计值与真实值之间的差异依概率收敛于 0。需要的正则性条件是分布不是肥尾的，以及条

件协方差矩阵随时间平滑变动。利用高频数据，错误设定的 ARCH 模型可以得到波动率的精确估计。在某种程度上，将变量平方值作平均的 GARCH 模型可解释为条件方差在时间 t 的非参数估计。离散时间模型也可以用连续时间扩散模型来近似。不同的 ARCH 模型通常有不同的扩散极限。正如 Nelson(1990b)所证明的那样，当离散时间模型出现难以处理的分布形式时，连续极限可能方便地产生近似预测值和其他矩的近似估计值。

Nelson 和 Foster（1994）研究了如何选择 ARCH 过程能一致、有效地估计生成数据的扩散过程的条件方差。他们得出了近似使用 ARCH 滤子所造成的测量误差的近似分布，其结果可用于比较不同的 ARCH 滤子的有效性，并且可以刻画渐进最优 ARCH 条件方差估计量的特性。他们推导了三个扩散模型的最佳 ARCH 滤子，并验证了许多 GARCH 模型的滤子性质。例如，如果数据生成过程是由具有独立布朗运动（ $\rho=0$ ）且 $\delta=1$ 的扩散方程（3.4）给出的， σ_t^2 的渐进最优滤子方程就设定了 y_t 的漂移 $\hat{\mu}=\mu$ 以及条件方差

$$\hat{\sigma}_{t+h}^2 = w.h + (1-\phi h-\alpha h^{1/2})\hat{\sigma}_t^2 + h^{1/2}\alpha\varepsilon_{y,t+h}^2 \qquad (4.1)$$

其中 $\varepsilon_{y,t+h} = h^{-1/2}[y_{t+h}-y_t-E_t(y_{t+h}-y_t)], w=\varpi\phi 且 \alpha=\xi/\sqrt{2}$ 。

所以，带有独立布朗运动的（3.4）的渐进最优滤子就是 GARCH（1，1）模型。当 w_y 和 w_σ 相关时，（4.1）的 GARCH（1，1）模型就不再是最优的。Nelson 和 Foster（1994）证明：Engle 和 Ng（1993）提出的非线性非对称 GARCH 模型在这一情形中满足最优化条件。Nelson 和 Foster（1994）还研究了当数据是由一个近似扩散的离散时间序列生成时，各种 ARCH 滤子的特征。他们的发现对于在实证研究中选择 ARCH 滤子的函数形式具有重要意义。

连续渐近记录极大地增强了我们对于当抽样频率增加时，连续时间随机微分方程与离散时间 ARCH 模型之间关系的理解。类似的，对随时间变化的波动率建模，尤其是在选择应用较高频的观测值还是较低频的样本观测值时，时频归并（temporal aggregation）起了重要作用。

通过高频数据可以得到更有效的参数估计。在有些情况下，研究者可能关注只可得到低频观测值时的高频模型的参数估计。

Diebold（1988）提出了时频归并问题。他证明，当抽样频率降低时条件异方差在极限意义上消失，而在流量变量（flow variable）的情形下，低频观测值的边际分布趋于正态分布。

Drost 和 Nijman（1993）研究了对股票或流量变量建模时，GARCH 类过程在时频归并的情况下是否闭合。

如果进行一些限制，就可以回答这个问题。这里采用了三种定义的 GARCH：如果可以选择 α_0、$\alpha_i, i=1,2,\cdots,q 和 \beta_i, i=1,2,\cdots,p$ ，使得 $\varepsilon_t = y_t h_t^{-1/2}$ 是均值为 0、方差为 1 的 i.i.d.，那么（2.2）中的变量 y_t 序列由一个强式 GARCH 过程生成；如果 $E[y_t \mid y_{t-1}, y_{t-2},\cdots]=0 且 E[y_t^2 \mid y_{t-1}, y_{t-2},\cdots]=h_t$ ，则 y_t 序列就称为半强式 GARCH；而当 $P[y_t \mid y_{t-1}, y_{t-2},\cdots]=0 且 P[y_t^2 \mid y_{t-1}, y_{t-2},\cdots]=h_t$ ，其中 P 代表根据常数项、$y_{t-1}, y_{t-2},\cdots, y_{t-1}^2, y_{t-2}^2,\cdots$ 求得的最优线性预测，那么 y_t 序列就是弱式 GARCH（p,q）。

Drost 和 Nijman（1993）的最大发现是股票或流量变量的对称弱式 GARCH 类过程在时频归

并条件下是闭合的。这意味着如果高频过程是对称（弱）GARCH，低频过程也是对称弱 GARCH。低频过程条件方差的参数取决于对应的高频过程的均值、方差和峰度。对于具有 $\sum_{i=1}^{q}\alpha_i+\sum_{i=1}^{p}\beta_i<1$ 的 GARCH 过程，随着抽样频率的增加，条件异方差性也消失了。强式或半强式 GARCH 类过程在时频归并条件下一般不是闭合的。这意味着如果观测值频率与数据生成的频率不是恰好对应时，强式或半强式 GARCH 过程只接近于数据生成过程。

Drost 和 Werker（1995）在随后的论文中，研究了连续时间 GARCH 过程，即增量 $X_{t+h}-X_t$，$t\in hN$，对于每一个固定的时间间隔 h>0 都是弱式 GARCH 过程的性质。显然，根据 Drost 和 Nijman（1993）的结论，连续时间 GARCH 过程不可能是强式或半强式 GARCH，因为这类过程在时频归并条件下不是闭合的。

潜在连续性时间 GARCH 过程的假设导致了相关联的离散 GARCH 模型的峰度大于 3，这意味着它是肥尾的。Drost 和 Werker（1995）说明了如何识别连续时间扩散过程的参数与离散时间 GARCH 的参数。连续时间模型和离散时间模型参数之间的关系可用于以相当直接的方式从离散时间观测值估计扩散模型。

Nijman 和 Sentana（1993），通过证明独立单变量 GARCH 过程的同期加总（contemporaneous aggregation）将会产生弱 GARCH 过程，补充了 Drost 和 Nijman（1993）的结论。接着他们通过证明多变量 GARCH 过程生成的变量的线性组合也是弱 GARCH，拓展了这一发现，同样，多变量 GARCH 模型的边际过程也是弱 GARCH。最后，作者通过模拟实验得出，在很多情况下，当样本容量扩大时，假设条件正态分布的强式 GARCH 过程收敛于接近弱式 GARCH 参数的值估计量便是 ML。

有关 GARCH 过程时频归并和同期加总的研究结论表明，GARCH 过程的线性变换通常只是弱式 GARCH。

4. 3. 波动率预测

建立时间序列模型通常是为了获得样本之外的预测值。许多作者已经研究了具有时间依赖的条件异方差模型的预测问题。对于具有 ARCH 误差和 GARCH 误差的时间序列模型，Engle 和 Kraft（1983）以及 Engle 和 Bollerslev（1986）分别得到了多步预测误差方差的表达式。Bollerslev（1986）、Granger、White 和 Kamstra（1989）都注意到方差随时间变化的向前一步预测区间的构建。Baillie 和 Bollerslev（1992）研究了一个带有 ARMA-GARCH 扰动项的单一方程回归模型，由此得出了最小 MSE 预测。他们还导出了扰动项为 GARCH（1，1）的动态模型预测误差分布的矩。这些矩在使用 Cornish-Fisher 渐近扩展式构建预测区间时会有用处。Geweke（1989）在 Bayesian 的背景下通过数值积分，得到扰动项为 ARCH 的线性模型的多步向前预测误差密度。

Nelson 和 Foster（1995）推导了当观测数据为高频时，错误设定的 ARCH 对时间序列过程以及波动率的预测有良好表现的条件。与 Nelson 和 Foster 在 1994 年得到的成功过滤的条件一致，基本要求是 ARCH 模型正确设定所有状态变量前二阶矩函数形式。

为了说明如何构造预测误差方差的估计，考虑平稳 AR（1）过程：

$$y_t=\phi y_{t-1}+u_t \quad , \qquad (4.2)$$

其中，$u_t=\varepsilon_t h_t^{1/2}$ 是一个（2.2）的 GARCH（1，1）过程，在 t 时 y_{t+s} 的最小 MSE 预测是 $E_t(y_{t+s})=\phi^s y_t$，预测误差 $w_{ts}=y_{t+s}-\phi^s y_t$ 可表示为 $w_{ts}=u_{t+s}+\phi u_{t+s-1}+\cdots+\phi^{s-1}u_{t+1}$。它在 t 时的条件方差为

$$Var(w_{ts}) = \sum_{i=0}^{s-1} \phi^{2i} E_t(u_{t+s-i}^2), s > 0 \qquad (4.3)$$

可以通过递归计算得出。从 u_t 的 GARCH（1，1）过程导出 u_t^2 的 ARMA 表达式[参见 Bollerslev(1986)]：

$$u_t^2 = \alpha_0 + (\alpha_1 + \beta_1)u_{t-1}^2 - \beta_1 v_{t-1} + v_t, \qquad (4.4)$$

其中 $v_t = u_t^2 - h_t$。正如 Engle 和 Bollerslev（1986）所证明的那样，从（4.4）的表达式易于得到(4.3)的 r.h.s.期望值

$$E_t(h_{t+s}) = E_t(u_{t+s}^2) = \alpha_0 + (\alpha_1 + \beta_1)E_t(u_{t+s-1}^2), s > 1 \quad , \qquad (4.5)$$

当预测的时间跨度扩大时，最优预测单调收敛于无条件方差 $\alpha_0 /(1 - \alpha_1 - \beta_1)$。对于 IGARCH（1，1）模型，对条件方差的冲击是持续的，$E_t(h_{t+s}) = \alpha_0(s-1) + h_t$。表达（4.5）可作为未来波动率。Baillie 和 Bollerslev（1992）把 $E_t(h_{t+s})$ 的条件 MSE 表达为 t+s 时条件方差的预测值。

5. 结论

在这一章，我们对使用 GARCH 过程对随时间变化的波动率建模的有关文献进行了综述。在回顾大量的成果时我们更强调近期的发展。

自 1982 年 Engle 发表了突破性的论文之后的不到 15 年时间里，对 GARCH 模型的理解以及它们在经济时间序列中的应用都取得了重大进展。这些进展极大地改变了时间序列实证研究所采用的方法。同时，时间序列的统计特性，特别是现有模型还无法解释的金融时间序列的统计特性，已经导致了对波动率建模的发展。偏度的发现及把偏斜相关（Skewed Correlation）定义为 $[(\sum_t y_t^2 y_{t+k})/(T\sigma^3\sqrt{k})]$ 促进了非对称 GARCH 模型的发展。GARCH 模型中条件正态分布新生超峰度的存在使学生 GARCH 模型和 GARCH 跳跃模型得到了应用。使用具有随机趋势成分的方差成分模型，条件方差的连续性也已经被模型化。

条件协方差和相关系数中的时间变差的发现导致了多变量 GARCH 模型和因子—GARCH 模型的发展。因子—GARCH 模型有许多吸引人的特点。第一，易于根据经济理论对它们作出解释（因子模型中的套利定价理论在金融中已得到广泛应用）；第二，它们允许在变量的高维向量中，随时间变化的方差和协方差有更简约的表达式；第三，它们既可以解释可观测因子又可以解释不可观测因子；第四，变量的共同特征有令人感兴趣的含义，这些共同特征可以用直接的方法来检验；第五，它们在许多例子中都拟合得很好。

随时间变化的波动率的函数形式已经吸引了许多研究者的注意，它达到这种程度，以致于有人疑惑：设计新的 GARCH 表达式的收益是否仍然为正。这一点在第 2 节中很明显。当模型即使不能完全相互替代，有些也可以非常相似时，Nelson 和 Foster 利用 GARCH 作为滤子来估计潜在扩散模型的条件方差，这一结果为选择 GARCH 模型的函数形式提供了新的视角。在给定的扩散过程中，一些 GARCH 模型将是最优（最有效）滤子，而其它有着类似特征的模型却可能不是

最优。Nelson 和 Foster（1994）研究表明，有关潜在扩散过程形式的先验知识在选择 GARCH 模型的函数形式时是有用的。

正如 Anderson（1992 和 1994）所证明的那样，GARCH 过程属于确定性、条件异方差类的波动过程。GARCH 似然函数易于求值，GARCH 设定具有适应波动率随时间变化的能力，尤其是它产生在许多时间序列中已经发现的平方值相关的灵活、简约的表达式（与用 ARMA 方法得到的条件均值的简约表达式相比），这些特征都使 GARCH 模型得到了广泛应用。随机波动率模型的历史是短暂的，但它是除了 GARCH 模型之外另一可供选择的参数简约方法。GARCH 模型的吸引人的特点之一是在拟合许多时间序列的波动率的时间变差时，只需要很少的参数，基于似然推断的随机波动率模型却需要数值积分或要使用 Kalman 滤子。正如第 3 节所提及的，这些问题许多现在已经解决。GARCH 模型与随机波动率模型的统计特征是不同的。在金融时间序列的基础上比较这些模型[参见 Danielson(1994)，Hsieh(1991)以及 Jacquier 等(1995)和 Ruiz(1993)]，可以得出结论，这些模型对各种矩函数提出了不同权重。对这些模型的选择通常是一个实证问题。

在其他情形中，GARCH 模型更优，因它给出了潜在扩散模型方差的最优滤子。当必需设定潜在因子条件时，具有不可观测因子的因子—GARCH 模型将导致随机波动率成分。人们认为这两类波动率模型的分界不再是那么明显。

GARCH 过程时频归并的结果表明弱式 GARCH 是更为普遍的情形。由于加总的原因，依赖于强式 GARCH 的模型充其量是数据生成过程的最好近似。从实用的观点来看，用数据信息选择模型可能是最合适的。

将来研究的主题是增进我们对不同序列和不同市场波动率的理解，并在对它们之间的关系建模方面获得进一步的发展。多变量 GARCH、因子—GARCH 和随机波动率模型将会得到应用和推广。波动持续性的性质及其它从一个序列到另一序列的传递，波动持续性到条件期望收益的传递等问题，在将来会得到更大的关注。最后，检验和估计波动率模型及预测波动率的统计方法将很快提到研究日程上来。特别地，在对经济时间序列条件分布的时间变差建模方面，非参数和半参数方法展现了新的视野。

参考文献

Anderson, T. G. (1992). Volatility. Department of Finance, Working Paper No. 144, Northwestern University.

Anderson, T. (1994). Stochastic autoregressive volatility: A framework for volatility modeling. Math. Finance 4, 75-102.

Baillie, R. T. and T. Bollerslev (1990). A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange rate markets. J. Internat. Money Finance 9, 309-324,

Baillie, R. T. and T. Bollerslev (1992). Prediction in dynamic models with time-dependent conditional variances. J. Econometrics 52, 91-113.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1993). Fractionally integrated generalized auto-regressive conditional heteroskedasticity. Michigan State University, Working Paper.

Baillie, R. T. (1994) Long memory processes and fractional integration in econometrics. Michigan State University, Working Paper.

Ball, C. A. and A. Roma (1993). A jump diffusion model for the European Monetary System. J. Internat. Money Finance 12, 475-492.

Ball, C. A. and W. N. Torous (1985). On jumps in common stock prices and their impact on call option pricing. J. Finance 40, 155-173.

Bera, A, K. and S. Lee (1990). On the formulation of a general structure for conditional hetero-skedasticity. University of Illinois at Urbana-Champaign, Working Paper.

Bera, A. K., S. Lee, and M. L. Higgins (1992). Interaction between autocorrelation and conditional heteroskedasticity: A random coefficient approach. J. Business Econom. Statist. 10, 133-142.

Bera, A. K. and S. Lee (1993). Information matrix test, parameter heterogeneity and ARCH. Rev. Econom. Stud. 60, 229-240.

Bera, A. K. and M. L. Higgins (1995). On ARCH models: Properties, estimation and testing. In: Oxley L., D. A. R.George, Roberts, C. J., and S. Sayer eds., Surveys in Econometrics, Oxford, Basil Blackwell, 215-272.

Black, F. (1976). Studies in stock price volatility changes. Proc. Amer. Statist. Assoc., Business and Economic Statistics Section 177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31, 307-327.

Bollerslev, T., R. F. Engle, and J. M, Wooldridge (1988). A capital asset pricing model with time varying covariances. J. Politic. Econom. 96, 116-131.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). ARCH modeling in finance: A review of the theory and empirical evidence. J. Econometrics 52, 5-59.

Bollerslev, T. and J. M. Wooldridge (1992). Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances. Econometric Rev. 11, 143-172.

Bollerslev, T. and I. Domowitz (1993). Trading patterns and the behavior of prices in the interbank foreign exchange market, J. Finance, to appear.

Bollerslev, T. and R. F. Engle (1993). Common persistence in conditional variances, Econometrics 61, 166-187.

Bollerslev, T. and H. O. Mikkelsen (1993). Modeling and pricing long-memory in stock market volatility. Kellogg School of Management, Northwestern University, Working Paper No. 134.

Bollerslev, T., R. F. Engle and D. B. Nelson (1994). ARCH models. Northwestern University, Working Paper, prepared for The Handbook of Econometrics Vol. 4.

Bougerol, Ph. and N. Picard (1992). Stationarity of GARCH processes and of some nonnegative time series. J. Econometrics 52, 115-128.

Brock, A. W., W. D. Dechert and J. A. Scheinkman (1987), A test for independence based on correlation dimension. Manuscript, Department of Economics, University of Wisconsin, Madison.

Brock, A.W., D. A. Hsieh and B. LeBaron (1991). Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence. MIT Press, Cambridge, MA.

Cai, J. (1994). A Markov model of switching-regime ARCH. J. Business Econom. Statist. 12, 309-316.

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using GARCH. J. Appl. Econometrics 3, 279-294.

Crouhy, M. and C. M. Rockinger (1994). Volatility clustering, asymmetry and hysteresis in stock returns: International evidence. Paris, HEC-School of Management, Working Paper.

Crowder, M. J. (1976). Maximum likelihood estimation with dependent observations, J. Roy. Statist. Soc. Ser. B 38, 45-53.

Danielson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. J. Econometrics 64, 375-400.

Davidian, M. and R. J. Carroll (1987). Variance function estimation. J. Amer. Statist. Assoc. 82, 1079-1091.

Demos A. and E. Sentana (1991). Testing for GARCH effects: A one-sided approach London School of Economics, Working Paper.

De Vries, C. G. (1991). On the relation between GARCH and stable processes. J. Econometrics 48, 313-324.

Diebold, F. X. (1987) Testing for correlation in the presence of ARCH. Proceedings from the ASA Business and Economic Statistics Section, 323-328.

Diebold, F. X. (1988). Empirical Modeling of Exchange Rates. Berlin, Springer-Verlag,

Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: A multivariate latent factor ARCH model. J. Appl. Econometrics 4, 1-21.

Diebold, F. X. and J. A. Lopez (1994). ARCH models. Paper prepared for Hoover K. ed., Macro-econometrics: Developments, Tensions and Prospects.

Ding, Z., R. F Engle, and C. W. J. Granger (1993). A long memory property of stock markets returns and a new model. J. Empirical Finance 1, 83-106.

Drost F. C. and T. E. Nijman (1993). Temporal aggregation of GARCH processes. Econometrica 61, 909-927.

Drost, F. C. and B. J. M. Werker (1995). Closing the GARCH gap: Continuous time GARCH modeling. Tilburg University, paper to appear in J. Econometrics.

Engel, C. and J. D. Hamilton (1990). Long swings in the exchange rate: Are they in the data and do markets know it ? Amer. Econom. Rev. 80, 689--713.

Engle, R. F. (I982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica 50, 987-1008.

Engle, R. F. and D. F. Kraft (1983). Multiperiod forecast error variances of inflation estimated from ARCH models. In: Zellner, A. ed., Applied Time Series Analysis of Economic Data, Bureau of the Census, Washington D.C., 293-302.

Engle, R. F. and T. Bollerslev (1986). Modeling the persistence of conditional variances. Econometric Rev. 5, 1-50.

Engle, R. F., D. M. Lilien, and R, P. Robins (1987). Estimating time varying risk premia in the term structure: The ARCH-M model, Econometrics 55, 391-407.

Engle, R. F. (1990). Discussion: Stock market volatility and the crash of 87. Rev. Financ. Stud. 3, 103-106.

Engle, R. F., V. K. Ng, and M. Rothschild (1990). Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills. J. Econometrics 45, 213-238,

Engle, R. F. and G. Gonzalez-Rivera (1991). Semiparametric ARCH models. J. Business Econom. Statist. 9, 345-359.

Engle, R. F. and V. K. Ng (1993). Measuring and testing the impact of news on volatility, J. Finance 48, 1749-1778.

Engle, R. F and G. G. J. Lee (1993): Long run volatility forecasting for individual stocks in a one factor model. Unpublished manuscript, Department of Economics, UCSD.

Engle, R. F. and S. Kozicki (1993). Testing for common features (with discussion). J. Business Econom. Statist. 11, 369-380.

Engle, R. F. and R. Susmel (I993). Common volatility and international equity markets. J. Business Econom. Statist. 11, 167-176.

Engle, R. F. and G. G. J. Lee (1994). Estimating diffusion models of stochastic volatility. Mimeo, University of California at San Diego.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. Econometric Theory 11, 122-150.

French, K. R., G. W. Schwert and R. F. Stambaugh (1987). Expected stock returns and volatility. J. Financ. Econom. 19, 3-30.

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form. J. Econometrics 15, 211-244.

Gallant, A. R. and G. Tanchen (1989). Seminonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications. Econometrica 57, 1091-1120.

Gallant, A. R., D. Hsieh and G. Tauchen (1994). Estimation of stochastic volatility model with suggestive diagnostics. Duke University, Working Paper.

Geweke, J. (1989). Exact predictive densities for linear models with ARCH disturbances. J. Econometrics 40, 63-86.

Geweke, J. (1994). Bayesian comparison of econometric models. Federal Reserve Bank of Minneapolis, Working Paper.

Ghysels, E., A. C. Harvey and E. Renault (1995). Stochastic volatility. Prepared for handbook of Statistics, Vol. 14.

Glosten, L. R., R. Jagannathan, and D. Runkle (1993). Relationship between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48, 1779-1801.

Gouriéroux, C. and A. Monfort (1992). Qualitative threshold ARCH models. J. Econometrics 52, 159-199.

Gouriéroux, C. (1992). Modèles ARCH et Application Financières. Paris, Economica.

Gouriéroux, C., A. Monfort and E. Renault (1993). Indirect inference. J. Appl. Econometrics 8, S85-Sl18.

Granger, C. W. J., H. White and M. Kamstra (1989). Interval forecasting: An analysis based upon ARCH-quantile estimators. J. Econometrics 40, 87-96.

Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. J. Econom, Dynamic Control 12, 385-423.

Hamilton, J. D. (1989). Analysis of time series subject to changes in regime. J. Econometrics 64, 307-333.

Hamilton, J. D. and R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. J. Econometrics 64, 307-333.

Harvey, A. C., E. Ruiz and E. Sentana (1992). Unobserved component time series models with ARCH disturbances..1. Econometrics 52, 129-158.

Hentschel, L. (1994). All in the family: Nesting symmetric and asymmetric GARCH models. Paper presented at the Econometric Society Winter Meeting, Washington D.C., to appear in J. Financ. Econom. 39, nr. 1.

Higgins, M. L. and A. K. Bera (1992). A class of nonlinear ARCH models, Internat. Econom. Rev. 33, 137-158.

Hsieh, D. A. (1989). Modeling heteroskedasticity in daily foreign exchange rates. J. Business Econom. Statist. 7, 307-317.

Hsieh, D. (1991). Chaos and nonlinear dynamics: Applications to financial markets. J. Finance 46, 1839-1877.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. J. Finance 42, 281-300.

Jacquier, E., N. G. Polson and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models. J. Business. Econom. Statist. 12, 371-389.

Jorion, P. (1988). On jump processes in foreign exchange and stock markets. Rev. Finan. Stud. 1, 427-445.

Kim, S. and N. Sheppard (1994). Stochastic volatility: Likelihood inference and comparison with ARCH models. Mimeo, Nuffield College, Oxford.

King, M., E. Sentana and S. Wadhwani (1994). Volatility links between national stock markets. Econometrica 62, 901-933.

Kodde, D. A. and F. C. Palm (1986). Wald criteria for jointly testing equality and inequality restrictions. Econometrica 54, 1243-1248.

Krengel, U. (1985). Ergodic Theorems. Walter de Gruyter, Berlin.

Lee, J. H. H. (1991). A Lagrange multiplier test for GARCH models. Econom. Lett. 37, 265-271.

Lee, J. H. H. and M. L. King (1993). A locally most mean powerful based score test for ARCH and GARCH regression disturbances. J. Business Econom. Statist. 11, 17-27.

Lee, S. W. and B. E. Hansen (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econometric Theory 10, 29-52.

Li, W. K. and T. K. Mak (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. J. Time Series Analysis 15, 627-636.

Lin, W. -L. (1992). Alternative estimators for factor GARCH models-A Monte Carlo comparison. J. Appl. Econometrics 7, 259-279.

Linton, O. (1993). Adaptive estimation in ARCH models. Econometric Theory 9, 539-569.

Lumsdaine, R. L. (1992). Asymptotic properties of the quasi-maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models. Unpublished manuscript, Department of Economics, Princeton University.

Lumsdaine, R. L. (1995). Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: A Monte Carlo investigation. J. Business Econom. Statist. 13, 1-10.

Melino, A. and S. Turnbull (1990). Pricing foreign currency options with stochastic volatility. J. Econometrics 45, 239-266.

Nelson, D. B. (1990a). Stationarity and persistence in the GARCH(1,1) model. Econometric Theory 6, 318-334.

Nelson, D. B. (1990b). ARCH models as diffusion approximations. J. Econometrics 45, 7-38.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347-370.

Nelson, D. B. (1992). Filtering and forecasting with misspecified ARCH models I. J. Econometrics 52, 61-90,

Nelson, D. B. and C. Q. Cao (1992). Inequality constraints in univariate GARCH models. J. Business Econom. Statist. 10, 229-235.

Nelson, D. B. and D . P. Foster (1994). Asymptotic filtering theory for univariate ARCH models. Econometrica 62, 1-41.

Nelson, D. B. and D. P. Foster (1995). Filtering and forecasting with misspecified ARCH models II - Making the right forecast with the wrong model. J. Econometrics 67, 303-335.

Ng, V., R. F. Engle, and M. Rothschild (I992). A multi-dynamic-factor modal for stock returns. J. Econometrics 52, 245-266.

Nieuwland, F. G. M. C., W. F. C. Verschoor, and C. C. P. Wolff (1991). EMS exchange rates. J. Internat. Financial Markets, Institutions and Money 2, 21-42.

Nijman, T. E. and F. C. Palm (1993). GARCH modelling of volatility: An introduction to theory and applications. In: De Zeeuw, A. J. ed,, Advanced Lectures in Quantitative Economics II, London, Academic Press, 153-183.

Nijman, T. E. and E. Sentana (1993). Marginalization and contemporaneous aggregation in multivariate GARCH processes. Tilburg University, CentER, Discussion Paper No. 9312, to appear in J. Econometrics.

Pagan, A. R. and A. Ullah (1988). The econometric analysis of models with risk terms. J Appl. Econometrics 3, 87-105.

Pagan, A. R. and G. W. Schwert (1990). Alternative models for conditional stock volatility. J. Econometrics 45, 267-290.

Pagan, A. R. and Y. S. Hong (1991). Nonparametric estimation and the risk premium. In: Barnet, W. A., J. Powell and G. Tauchen, eds., Nonpurametric and Semiparametric Methods in Econometrics and Statistics, Cambridge University Press, Cambridge.

Pagan, A. R. (1995). The econometrics of financial markets. ANU and the University of Rochester, Working Paper, to appear in the J. Empirical Finance.

Palm, F. C. and J. P. Urbain (1995). Common trends and transitory components of stock price volatility. University of Limburg, Working Paper.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. J. Business 53, 61-65.

Ruiz, E. (1993). Stochastic volatility versus autoregressive conditional heteroskedasticity. Universidad Carlos III de Madrid, Working Paper,

Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. J. Econometrics 47, 67-84.

Schwert, G. W. (1989). Why does stock market volatility change over time? J. Finance 44, 1115-1153.

Sentana, E. (1991). Quadratic ARCH models: A potential re-interpretation of ARCH models. Unpublished manuscript, London School of Economics.

Sentana, E. (1992). Identification of multivariate conditionally heteroskedastic factor models. London School of Economics, Working Paper.

Taylor, S. (1986). Modeling Financial Time Series. J. Wiley & Sons, New York, NY.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. Math. Finance 4, 183-204.

Tsay, R. S. (1987). Conditional heteroskedastic time series models. J. Amer. Statist. Assoc. 82, 590-604.

Vlaar, P. J. G. and F. C. Palm (1993). The message in weekly exchange rates in the European Monetary System: Mean reversion, conditional heteroskedasticity and jumps. J. Business. Econom. Statist. 11, 351-360.

Vlaar, P. J. G. and F. C. Palm (1994). Inflation differentials and excess returns in the European Monetary System. CEPR Working Paper Series of the Network in Financial Markets, London.

Weiss, A. A. (1986), Asymptotic theory for ARCH models: Estimation and testing. Econometric Theory 2, 107-131.

Zakoian, J. M. (1994). Threshold heteroskedastic models. J. Econom. Dynamic Control 18, 931-955.