

用 GARCH 模型与隐含波动率预测金融波动率

骆 桦,王 爽

(浙江理工大学理学院,杭州 310018)

摘 要: 利用港股期权的数据,研究在不同期限内的 GARCH 模型与隐含波动率的表现,并且用已实现波动率、M-Z 回归分析、误差项和损失函数对预测效果进行评估。研究结果表明:短期 GARCH 模型预测效果较好,长期隐含波动率预测效果较好;由于期权价格能够更加全面的反映市场信息,所以隐含波动率对于波动率的预测效果更好。

关键词: 港股期权;GARCH 模型;隐含波动率;已实现波动率

中图分类号: F830.91 **文献标志码:** A **文章编号:** 1673-3851 (2016) 02-0322-05 **引用页码:** 030804

0 引 言

波动率是描述投资风险以及金融市场变化情况的重要指标之一,它与金融市场的不稳定性密切相关。估计和预测波动率可以帮助投资者合理分配资产,提高资产收益率,降低投资风险。

在金融市场中,波动率的研究可以分为两大类:第一类是利用价格的历史数据预测将来的波动率,如 1982 年 Engle^[1]首次提出 ARCH 模型并用线性形式表出。以及由 Bollerslev^[2]提出的 GARCH 模型。第二类是利用隐含波动率来预测,即利用 Black-Scholes 公式^[3]反推出波动率。利用上述两类方法时,存在一大问题:模型是否能更好地拟合样本内的数据又能对样本外的数据作出精确的预测。朱雅宁^[4]以上海期货交易所 AU1406 日收益率为数据,利用损失函数对 GARCH、EGARCH、TGARCH 3 种模型波动率的预测结果进行比较,认为对称的 GARCH 模型预测效果比较好,为国内黄金生产企业在市场风险度量与预测以及投资是有帮助的。司训练等^[5]研究了上海证券日收益率,得出了能够描述波动率的最优模型,提高了资产收益以及投资的准确性。陈彦晖^[6]考虑均值回归效应、溢

出效应和周内效应后,认为 ARMA-GARCH 模型对恒指隐含波动率的预测能力有所提高,投资者的期权交易获得了较好的收益。

波动率具有波动集聚性,即大的波动后跟随着大的波动,小的波动后跟着小的波动。本文根据港股期权数据,研究在不同期限内的 GARCH 模型与隐含波动率的表现,探讨在不同市场环境下采用何种模型更为合理的问题。哪个模型对于短期的预测更为准确,哪个波动率对于长期的预测准确性更高。

1 隐含波动率与 GARCH 模型介绍

1.1 隐含波动率理论

Black 和 Scholes 在 1973 年给出了期权定价公式(BS 公式),下面简单介绍此公式的推导过程。

设 t 时刻的资产价格(如股票)为 $S(t)$,则 $S(t)$ 满足如下随机微分方程

$$dS = \mu S dt + \sigma S dB(t) \quad (1)$$

其中: μ 是期望回报率(漂移率), σ 为波动率, $B(t)$ 是标准布朗运动。

相应地,股票期权的价格为 $V(S, t)$,构造一个投资组合 Π ,买入一份看涨期权 V ,卖出 Δ 份股票,即:

$$\Pi = V - \Delta S,$$

在时间 $(t, t + dt)$ 内,此投资组合是无风险的。设无风险利率为 r ,且 r 为常数,则方程为

$$\Pi_{t+dt} - \Pi_t = r \Pi_t dt \quad (2)$$

由(1)、(2)和伊藤公式可得到著名的 Black-Scholes 方程:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \quad (3)$$

解此方程得期权定价公式:

$$V(S, t) = SN(d_1) - Ke^{-r(T-t)} N(d_2) \quad (4)$$

$$\text{其中: } d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, d_2 = d_1 -$$

$\sigma\sqrt{T-t}$, S 为 t 时刻的股票价格, T 为到期日, K 为执行价, $N(t)$ 是标准正态分布。

在式(4)中取定 S, K, T, r 的值,以及市场交易中市场期权的价格 V ,代入公式(4)中,反推出波动率 σ 的值,这一数值就是隐含波动率。

1.2 GARCH 模型

对于时间序列 $\{a_t\}$,若有:

$$\begin{cases} a_t = \sigma_t u_t, u_t \sim i. i. d. N(0, 1) \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i a_{t-i}^2 \end{cases} \quad (5)$$

则序列 $\{a_t\}$ 是一个 ARCH(q) 序列,模型(5)为 ARCH(q) 模型。其中: σ_t^2 为方差, q 为 ARCH(q) 模型的阶数, u_t 服从独立标准正态分布, $\alpha_0, \alpha_i \geq 0 \quad \forall i = 1, 2, \dots, q$ 。

在 ARCH(q) 模型的基础上, Bollerslev 提出如下的 GARCH(p, q) 模型:

$$\begin{cases} a_t = \sigma_t u_t, & u_t \sim i. i. d. N(0, 1) \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i a_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases} \quad (6)$$

其中: p, q 为阶数, $\alpha_0, \alpha_i, \beta_j$ 为参数且大于等于零 $\forall j = 1, 2, \dots, p, \forall i = 1, 2, \dots, q, u_t$ 服从独立标准正态分布。当 $p = q = 1$ 时,式(6)为

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (7)$$

这就是常用的 GARCH(1,1) 模型。

2 波动率模型的计算与预测效果的评价标准

2.1 波动率模型的计算

如何用 GARCH(1,1) 模型进行计算。在给定时刻 $t-1$ 的信息 F_{t-1} 条件下,由式(7)可以得到

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 a_t^2 + \beta_1 \sigma_t^2,$$

其中 a_t^2, σ_t^2 均已知。故向前一步预测可以表示为

$$\sigma_t^2(1) = \alpha_0 + \alpha_1 a_t^2 + \beta_1 \sigma_t^2 \quad (8)$$

由式(6)GARCH 模型的定义可知

$$a_t^2 = \sigma_t^2 u_t^2 \quad (9)$$

其中 $u_t \sim i. i. d. N(0, 1)$ 。由式(8)、(9)可得

$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2 + \alpha_1 \sigma_t^2 (u_t^2 - 1).$$

因为 $E(u_t^2 - 1 | F_{t-1}) = 0$, 则以 t 时刻为预测的原点向前两步预测表示为:

$$\sigma_t^2(2) = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2(1).$$

利用迭代的方法,最后可以得到向前 k 步预测公式,即:

$$\sigma_t^2(k) = \frac{\alpha_0 [1 - (\alpha_1 + \beta_1)^{k-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{k-1} \sigma_t^2(1)$$

只要 $\alpha_1 + \beta_1 < 1$, 那么当 $k \rightarrow \infty$ 时就有

$$\sigma_t^2(k) \rightarrow \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

所以由以上的推导过程可以计算出 GARCH(1,1) 模型的方差。进而也可以利用期权定价公式求出看涨期权的理论价格。

2.2 预测效果的评价标准

2.2.1 已实现波动率的计算

高频数据一般是指时间间隔较短的证券交易数据,例如股票 5 min 的价格走势图。近年来,对于高频数据的计算,已经推广到利用日内数据的对数收益率来估计当月的波动率。首先给出日对数收益率的定义。

定义:日对数收益率为日收盘价的一阶自然对数的差分,计算公式为:

$$r_t = \ln(S_t / S_{t-1}),$$

其中: r_t 为第 t 天的日对数收益率, S_t 为第 t 天的股票收盘价。假设 r_t 第 t 个月的对数收益率,第 t 个月的交易日为 m 天,第 t 个月的日对数收益率记为 $\{r_{t,i}\}_{i=1}^m$,则有:

$$r_t = \sum_{i=1}^m r_{t,i}.$$

假设条件方差与协方差均存在且 $\{r_{t,i}\}_{i=1}^m$ 为白噪声序列,则:

$$\text{Var}(r_t | F_{t-1}) = \sum_{i=1}^m \text{Var}(r_{t,i} | F_{t-1}) + 2$$

$$\sum_{i < j} \text{Cov}[(r_{t,i}, r_{t,j} | F_{t-1})] = n \text{Var}(r_{t,1}),$$

其中:第 $t-1$ 个月的已知信息为 F_{t-1} ,且 $\text{Var}(r_{t,1})$ 可通过日收益率求出,即:

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^m (r_{t,i} - \bar{r}_t)^2}{m-1},$$

其中, $\bar{r}_t = \frac{1}{m} \sum_{i=1}^m r_{t,i}$, 如果进一步要求 \bar{r}_t 均值为零, 则月收益率为:

$$\hat{\sigma}^2 = \frac{m \sum_{i=1}^m (r_{t,i} - \bar{r}_t)^2}{m-1} \approx \sum_{i=1}^m r_{t,i}^2 \quad (10)$$

2.2.2 M-Z 回归分析

M-Z 回归方程是真实波动率 $\hat{\sigma}^2$ 与预测波动率 $\tilde{\sigma}^2$ 建立的一元线性回归模型, 即:

$$\hat{\sigma}^2 = \alpha + \beta \tilde{\sigma}^2 + \epsilon_t,$$

其中: ϵ_t 为误差干扰项, 真实波动率 $\hat{\sigma}^2$ 是未知的, 可用已实现波动率代替计算。若模型的选择是准确的, 则 $E(\hat{\sigma}^2) = \tilde{\sigma}^2$, 那么要求有 $\alpha = 0, \beta = 1$ 。进一步, 我们可以利用误差项 ϵ_t 值的大小, 对实验结果进行检验。

2.2.3 损失函数

对于波动率的研究, 误差问题是无法避免的, 那么采用何种方式对估计结果进行检验? 以及按照何种标准来选择模型? 模型的检验标准有很多, 但是希望能够找到一种计算简单又可以包含有用信息的检验标准。所以本文采用 3 种误差指标对预测的结果进行评估, 分别是: V_{MAE} 、 V_{MAPE} 、 V_{RMSE} 。其中, V_{MAE} 是平均绝对误差, V_{MAPE} 是平均绝对误差百分比, V_{RMSE} 是均方根误差。3 种误差指标不仅能很好地反映出计算的精确度, 而且能判断出哪种模型的拟合效果更好。

主要思路是通过模型预测波动率与实际波动率之间产生的一些误差来进行评估。所得到的结果越小, 说明误差越小, 模型的估计结果就越准确。由于实际的波动率是未知的, 所以用收益率近似等于实际的波动率来进行比较。

3 个误差指标公式如下所示

$$V_{MAE} = \frac{1}{n} \sum_{t \in S} |M_{F,t} - M_{R,t}|,$$

$$V_{MAPE} = \frac{1}{n} \sum_{t \in S} \frac{|M_{F,t} - M_{R,t}|}{M_{R,t}},$$

$$V_{RMSE} = \sqrt{\frac{1}{n-1} \sum_{t \in S} (M_{F,t} - M_{R,t})^2}.$$

其中: $M_{F,t}$ 是模型预测出的波动率的数值, $M_{R,t}$ 是计算出的实际波动率的数值, n 为样本数量。

3 数据的选取与计算

3.1 数据来源

探究隐含波动率必须选用合适的的数据, 虽然国内从 2015 年 2 月 9 号起已经发行了 50 ETF, 但是由于上市的时间短, 并且为指数期权, 作为研究对象还不够成熟, 所以本文选用港股期权。采用中国银行自 2009 年 1 月 1 日到 2015 年 2 月 27 日共 1515 个收盘价数据进行分析(除去休市日), 并用股票 2015 年 3 月的收盘价数据作为预测分析, 看涨期权中行法巴七四购 A 收盘价数据进行计算。基本信息如下表所示。

表 1 期权信息表

发行人	发行日期	最后交易日	行使价	形式	性质	发行价
法巴	2014-01-15	2015-04-21	3.70	欧式权证	认购权证	0.26

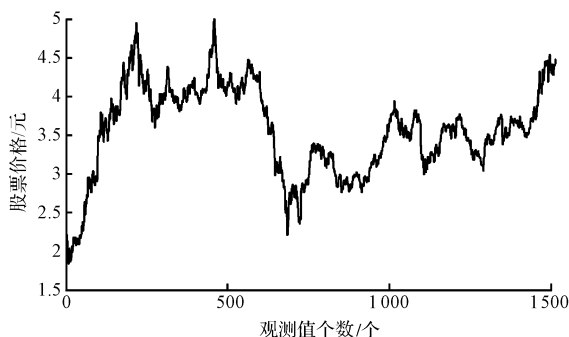


图 1 股票收盘价序列图

图 1 描述了股票收盘价时序, 股价的走势有明显的上升趋势, 为非平稳时间序列。在图 2 中, 对股

价进行一阶差分后, 序列基本平稳, 进而可用时间序列模型进行拟合与预测。收益率序列在大的波动后持续一段大的波动, 而小的波动后持续一段小的波动, 说明收益率的波动性存在聚集现象。

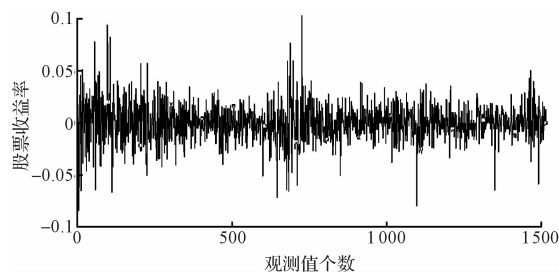


图 2 日收益率序列图

3.2 收益率序列的基本统计特征分析

利用 SAS 软件计算,得到收益率的统计特性,如表 2 所示。其中,J-B 是正态性检验的统计量。收益率序列 J-B 统计量的值相应的概率值为 0,说明收益率序列并不服从正态分布。表中的偏度与峰度值表明,收益率序列具有“尖峰厚尾”分布特征。建立模型前,仍需对收益率序列进行平稳性检验。

表 2 收益率的描述性统计表

均值	中位数	偏度	峰度	J-B 统计量	J-B 概率
0.000640	0.00000	0.108558	6.162283	633.8090	0.00000

3.3 平稳性检验(ADF 检验)

ADF 检验实质为假设检验。原假设 H_0 : 收益率序列存在单位根,备择假设 H_1 : 收益率序列不存在单位根(即平稳时间序列)。

表 3 收益率 ADF 检验

t 统计量	1% level	5% level	10% level	t 统计量概率
-38.26232	-3.434465	-2.863245	-2.567726	0.0000

表 3 中给出在检验水平分别为 1%、5%、10% 下 t 统计量的临界值,利用 t 分布假设检验, t 统计量为 -38.26232,均小于各个水平下 t 统计量的临界值,且相应的概率值为 0,故拒绝原假设,认为收益率序列不存在单位根,即为平稳时间序列。在平稳序列的基础上,便可以建立 ARCH 模型。

3.4 ARCH 模型与残差图

对收益率序列进行滞后 24 阶的自相关与偏自相关检验,可得到下列的均值方程,表达式为:

$$y_t = 2.62 \times 10^{-10} + u_t, u_t \sim i. i. d. N(0, \sigma_t^2).$$

如图 3 所示,大的波动后紧跟着一段大的波动,而小的波动后又紧跟着一段小的波动,说明序列存在 ARCH 效应。虽然在日收益率序列的残差图中可以发现,残差值与真实值之间非常相似,但是仍然要对均值方程进行 ARCH 检验。

表 6 GARCH(1,1) 模型的检验结果

	ARCH(0)	ARCH(1)	GARCH(1)	MSE	V_{MAE}	V_{MAPE}	V_{RMSE}
GARCH(1,1)	4.05×10^{-6}	0.039874	0.944869	0.00381	0.0130856	90.954047	0.0180277

如表 6 所示,根据相应的数据也可以写出 GARCH(1,1) 模型的表达式,即

$$\begin{cases} a_t = \sigma_t u_t, u_t \sim i. i. d. N(0, 0.00381) \\ \sigma_t^2 = 4.05E-06 + 0.039874a_t^2 + 0.944869\sigma_{t-1}^2 \end{cases}$$

3.6 GARCH(1,1) 的 ARCH 效应检验

为检验对序列的模型构建是否合理,还需要对残差序列进行 ARCH 效应检验。

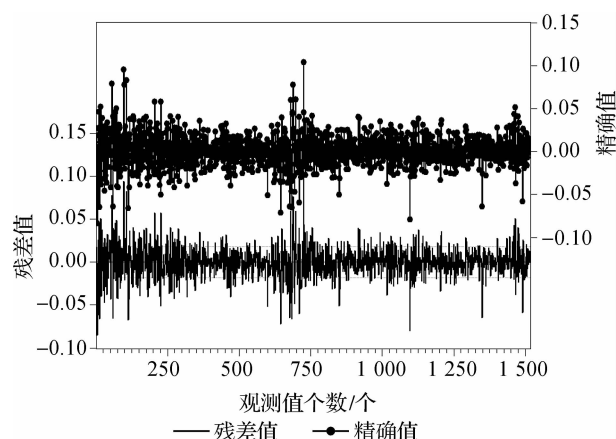


图 3 日收益率序列残差

原假设 H_0 : 残差序列不存在 ARCH 效应,备择假设 H_1 : 残差序列存在 ARCH 效应。

表 4 ARCH 效应检验

F 统计量	F 统计量 概率值	$T \times R^2$ 统计量	$T \times R^2$ 统计 量概率值
1.477593	0.1709	10.37522	0.1683

利用 ARCH-LM 检验可以得到 F 统计量的值和 $T \times R^2$ 的值相对应的概率,其差异不显著,所以拒绝原假设,认为收益率序列具有 ARCH 效应,可以建立 GARCH 模型。

3.5 建立 GARCH 模型

在 GARCH 模型中,考虑到模型的稳健性,所以只比较以下 4 种模型:

表 5 GARCH 模型比较

	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)	GARCH(2,2)
AIC	-5.361747	-5.360680	-5.360509	-5.361674
BIC	-5.351206	-5.346625	-5.346455	-5.344106

利用最小信息准则 AIC 与 BIC,具体数值如表 5 所示,则建立 GARCH(1,1) 模型为最优模型。

表 7 GARCH(1,1) 模型的 ARCH 效应检验

F 统计量	F 统计量 概率值	$T \times R^2$ 统计量	$T \times R^2$ 统计 量概率值
0.354530	0.9283	2.490827	0.9278

从表 7 中得知,ARCH 效应检验的 F 统计量与 $T \times R^2$ 统计量相应的概率显著,所以可以接受原假设,即残差序列不存在 ARCH 效应,也说明序列中

的信息被进行了有效的提取,序列没有统计规律,因而停止对序列的分析,所以 GARCH(1,1) 模型拟合效果好。

3.7 GARCH 模型与隐含波动率预期结果比较

对 GARCH 模型于隐含波动率预测效果进行比较,如表 8 所示。

表 8 两种模型的预测效果评估

同期	模型	损失函数			M-Z 回归		
		V_{MAE}	V_{MAPE}	V_{RMSE}	α	β	ϵ_t
以 1 个星期	GARCH(1,1)	0.02445	0.0369	103.1815	-0.0004	0.5663	-7.408×10^{-5}
为期限	隐含波动率	0.0283	0.0416	104.7476	0.0002	-0.0006	-0.0025
以 1 个月	GARCH(1,1)	0.0619	0.0822	102.8757	0.0027	-0.1633	-0.0008
为期限	隐含波动率	0.0247	0.0204	99.3896	-0.0061	0.6916	-0.11345

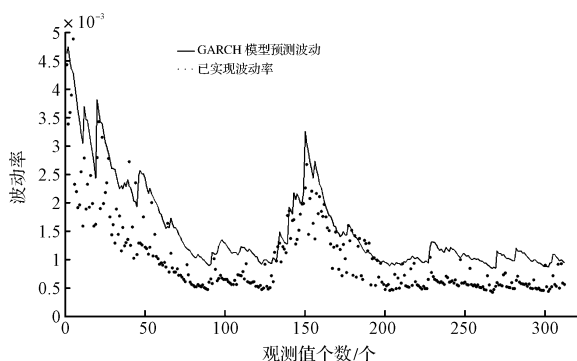


图 4 已实现波动率与 GARCH 模型预测波动率比较

本文利用损失函数、M-Z 回归分析和误差项 ϵ_t , 得到了相同的检验结果,即在以 1 周为期限进行预测时,GARCH 模型的预测结果优于隐含波动率模型;在以一个月为期限进行预测时,隐含波动率模型的预测结果优于 GARCH(1,1) 模型。从图 4 中也可以看出,在以周为期限时,已实现波动率与 GARCH 模型预测的波动率比较相近。

4 结 论

本文主要对股票的收益率的波动率变化进行了分析,利用 SAS 软件计算出 GARCH 模型与隐含波动率模型对波动率预测的表现,研究结果表明:GARCH 模型是利用历史数据计算得出,更好地体现了历史信息的影响,但不同时期的经济环境和企业的经营状况不同,故利用 GARCH 模型计算出

的波动率只适合用于短期(如一周)的预测。隐含波动率是根据现有的期权价格反推出波动率,而期权价格已经反映了市场环境的变化和企业经营信息的对价格的影响,所以较长时期(如一个月)的预测,隐含波动率效果更好。

参考文献:

- [1] ENGLE R F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation [J]. *Econometrica*, 1982, 50: 987-1007.
- [2] BOLLERSLEV T. Generalized autoregressive conditional Heteroskedastic [J]. *Journal of Econometrics*, 1986, 31: 307-327.
- [3] BLACK F, SCHOLES M. The pricing of options and corporate liabilities [J]. *Journal of Political Economy*, 1973, 81(3): 637-654.
- [4] 朱雅宁. 基于中国黄金期货市场数据的 GARCH 族模型拟合评价[J]. *时代金融*, 2015(2): 12-13.
- [5] 司训练, 张旭峰, 高蒙. 基于 GARCH 模型的金融风险拟合分析[J]. *经济学*, 2014, 23(5): 12-17.
- [6] 陈彦晖. 基于 ARMA-GARCH 模型的恒指隐含波动率指数预测及其在期权交易中的应用[J]. *经济数学*, 2014, 31(4): 28-35.
- [7] 郑振龙, 黄慧舟. 波动率预测: GARCH 模型与隐含波动率[J]. *数量经济技术经济研究*, 2010, 1: 140-150.
- [8] 黄海南, 钟伟. GARCH 类模型波动率预测评价[J], 2007, 15(6): 13-19.

Using GARCH Model and Implied Volatility to Predict Financial Volatility

LUO Hua, WANG Shuang

(School of Science, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: This Research studies GARCH model and implied volatility within different durations with the data of Hong Kong stock, and assesses the predictive effect by utilizing realized volatility, M-Z regression analysis, error term and loss function. Study shows that: In the short term, the predictive effect of GARCH model is better; while in the long term, the implied volatility has good predictive effects. Because the option price can reflect more comprehensive market information, the predictive effect of implied is better.

Key words: Hong Kong stocks; GARCH model; implied volatility; realized volatility

(责任编辑: 张祖尧)