## 基于GARCH族模型的中日韩股指波动率研究

资产的风险是由波动率来度量的，被定义为资产收益序列的标准差。研究波动率可以提高对风险的认识，从而有效的为风险定价，提高金融市场配置资源的效率，降低经济运行的整体成本。因此，对波动率的研究逐渐成为金融学术界和投资者都极其重视的中心问题。真实的波动率无法直接从市场中观测得到，不能精确计算，所以只能借助于估算和预测模型。

对波动率的估算和预测主要分为两类：一类是历史信息法，通过过去预测将来，背后的逻辑是市场具有内在的稳定性和周期性。随机波动率和GARCH族时间序列方法是最常见的两种。另一类是隐含波动率，通过期权的价格倒推出市场对未来波动率的预期。期权作为一种转移和对冲风险的重要工具，其价格包含的信息非常丰富，蕴含了市场对未来波动率的预测。尤其是股指期权，因更具流动性和综合性，形成的价格包含了容量最大和最具前瞻性的预测信息，并且不断动态调整。因此股指期权和波动率预测更具紧密联系。

本文内容主要包括两部分：一、简述波动率的模型；二、以沪深300股票指数和国外部分股指为对象，对比GARCH族方法对波动率的建模。

历史波动率(Historical Volatility)

 20世纪70年代之前，同一资产的波动率是被假定为一个恒定常数，不考虑波动率时变的情况。1952年，经典的Markowitz投资组合分析中，开始以回报率的方差作为风险度量。采用无偏估计量，对波动率的估算如下：

 

 其中，是收益率，是收益率的均值，波动率的估计值的平方(方差)等于之前一段时间收益的方差的无偏估计量。波动率被假定为恒定，这个波动率估计值被直接视为对未来波动率的预测。这种方法是从历史数据以相同权重估计得来的波动率，所以被称为历史波动率。

 历史波动率模型极其简单，但误差很大。

隐含波动率 (Implied Volatility)

 期权定价的B-S-M模型中，标的资产的波动率作为影响期权价格的一个变量，被引入的期权价格的计算公式中。例如欧式看涨期权的价格公式：







 其中：K是期权的执行价格，T是期权的期限，是无风险利率，是标准正态分布的累积概率分布函数，是0时刻的股价，是期权的价格。

 以上公式中涉及六个变量，其中期权的价格和期限、无风险利率、标的资产的市场价格、期权的执行价格都是已知的，而波动率是唯一不能从市场观察的数据。将已知的五个数据代入，可以反解出标的资产的波动率。这种波动率是隐含在期权的价格信息当中，是将期权价格中包含的对未来波动率的预期的分离出来得到的，因此被称为隐含波动率。

通过交易形成的期权价格包含了对标的资产波动率的预测。波动率的规律也影响了资产的交易，比如外汇期权中的波动率与行权价的非线性形式-波动率微笑，恰恰是外汇交易员在实际交易中的依据之一。

隐含波动率的优点是：金融市场每日行程的价格反映了投资者从历史数据和最新资讯中获取各种信息后形成的预期，是最具信息容量和前瞻性的波动率预测。

存在的问题是有两点前提不能完全成立：一是要求市场参与者比较理性，否则期权会出现各种噪音；二是期权定价模型本身是完全正确没有缺陷的。

已实现波动率(Realized Volatility)

 Andersen 和Bollerslev在1998年引入了已实现波动率,基本思想是当期波动率等于过去更高频的波动率的积分。但实际中波动率的积分是不可观测的，所以要进行估计。一般的处理是：第t天的已实现波动率的估计表示为当天内的高频收益平方和，即：

 

 这是对实际波动率的一致估计，当数据足够高频和离散时，得到的结果可以认为是对波动率的准确估计。已实现波动率包含了日内交易时段的所有波动率的信息，也就是隐含波动率的真实值。其问题是：由于股票市场不是24小时的连续交易，因此，能观察和记录到的高频股价数据只能反映交易时段内市场波动情况，而不包含交易时段外的市场波动信息，这就是股票市场从收盘到第二天开盘的“Close to Open”波动率。因此为准确刻画全天的市场波动率大小，在应用中可以采取某种尺度参数对实现波动率进行尺度变换。

 高频数据的时间间隔并不是越小越好。在高流动性的市场中，用很高频率的数据估计波动率往往会因为微观摩擦产生较大误差。Torben等人对道琼斯股票作实证分析时，对比了不同时间间隔的高频数据对波动率的计算效果，得出较佳的时间间隔是15-20分钟。

已实现波动率通常做为一种基准，代替不可观测的潜在真实波动率，来评价其它波动率模型的预测结果。

随机波动率 (Stochastic Volatility)

 随机波动率假定资产收益的方差服从某种滞后的随机过程，即将波动率本身视为随机变量。波动率其满足的随机微分方程可以分解为离散化的表示形式。1986年Taylor引入的一个简单随机波动率模型中，假设收益的均值为0，方差的对数的均值为0，其随机波动率的离散形式为：

 

  

 其中，是资产的收益率，是正态分布。这种简单模型是对Hull于1987年提出的随机波动率离散过程的一种近似。随机波动率模型的优点在于更容易建立连续的随机过程，便于理论分析。缺点在于参数估计的困难，主要是似然函数的表达式不易获得，所以很难对它进行估计。这也是随机波动率缺乏如GARCH族模型一般吸引力的原因。近几年来，随着计量经济学的发展，随机波动性模型的参数估计取得了明显进展，已提出许多有效的估计方法并用于实践。有很多文献是关于GARCH族模型和随机波动率模型对比的工作。

条件异方差模型

传统的计量模型中，随机扰动项的方差被假定为常数。金融时间序列常呈现波动的聚集性，为刻画这种异方差性，1982年由恩格尔(Engle.R)提出了条件方差ARCH模型，并由其学生波勒斯列夫(Bollerslev)于1986年发展成为广义自回归条件异方差模型(GARCH)。条件方差模型被广泛的应用到金融领域的非线性时间序列建模，在波动率预测方面尤其如此。

GARCH（p,q）模型可以表示为：

均值方程 ：

条件异方差方程：

其中是时刻t-1之前的全部信息，独立同分布满足条件：。

GARCH－M模型

GARCH-M模型是对均值方程增加一项，表达式为



其中服从GARCH(p, q)模型。假设模型旨在解释一项金融资产的回报率，那么增加的原因是每个投资者都期望资产回报率是与风险度密切联系的，而条件方差代表了期望风险的大小。所以GARCH-M模型适合于描述那些期望回报与期望风险密切相关的金融资产。

对GARCH模型的重要改进是引入不对称性，如TGARCH和EGARCH。

TARCH模型

TARCH模型(Threshold ARCH)模型具有如下形式的条件方差



其中是一个名义变量



由于引入,股价上涨信息（）和下跌信息（）对条件方差的作用效果不同。上涨时，其影响可用系数代表，下跌时为。若，则说明信息作用是非对称的。而当时，认为存在杠杆效应。

EGARCH模型

EGARCH模型，即指数(Exponential)模型，由Nelson在1991年提出的，其目的是为了刻划条件方差对市场中正、负干扰的反应的非对称性。此时条件方差为延迟扰动项的反对称函数：





模型中条件方差采用了自然对数形式，意味着杠杆效应是指数型的。若，说明信息作用非对称；若时，杠杆效应显著。因此EGARCH模型可以很好的刻划金融市场中的非对称性。此外由于被表示成指数形式，因而对模型中的参数没有任何约束，这是EGARCH模型的一大优点。

条件方差模型易于参数估计，其似然函数相对简单。存在的问题：基于历史信息预测未来波动率存在局限性，另外可能出现过度拟合的情况。

基于GARCH族模型，以中日韩的代表性股指为对象，对波动率做实证研究，考虑对比股指波动率在以下几个方面的特点：

1. 波动率的聚集性：波动率在一些时段上较高，另一些时段上较低；
2. 波动率的平稳性和跳跃；
3. 波动率对利空和利好消息反应程度的差异，即杠杆效应；
4. 波动率的记忆性。

 为保证时间跨度的一致，以沪深300股指的基期为起点，样本数据选取2004年12月31日至2012年6月1日的股指收盘价。因各国节假日的不同，样本数据量有差异。其中，沪深300股指有1800个观测值，香港恒生指数有1831个观测值，韩国综合指数有1846个观测值，日经指数有1819个观测值。数据来源是wind资讯终端与Yahoo Finance。将收盘价做对数差分处理后，得到股指日收益率序列。通过观察收益率序列图,可看出各国的股指收益率在样本区间内均表现出波动的聚集性，即异方差性质，可以考虑用条件方差模型建模。此外，中国沪深300指数波动性最为显著，波动率的跳跃也最明显。而日本N225指数的波动性最小，波动率表现出平稳性的一面。韩国综合指数的波动率的聚集性比较明显。在2008年金融危机前后，各股指的波动率都伴随着股价变动出现大幅跳跃。

股指收益率序列图



沪深300 香港HSI



韩国KSII 日本N225

 对比以上四个股指收益率序列的基本统计量，如下表所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 样本数据 | 均值 | 标准差 | 偏度 | 峰度 | J-B统计量 |
| 沪深300 | 0.000538 | 0.019552 | -0.375332 | 5.538285 | 525.1868 |
| 香港HSI | 0.000145 | 0.017697 | 0.055072 | 11.08350 | 4986.047 |
| 韩国KSII | 0.000388 | 0.015187 | -0.572284 | 9.679878 | 3532.845 |
| 日本N225 | -0.000170 | 0.016314 | -0.558374 | 12.20694 | 6519.217 |

从表中数据可知，均值都很小，都接近为零。偏度系数显著不为零，四个股指都具有非对称性，都有偏斜。峰度系数均远大于3，J-B统计量都说明不服从正态分布。这说明股指收益率序列具有金融时间数据的特征，即尖峰厚尾。在通过ADF方法检验序列平稳性后，再通过自相关检验观察序列自相关的情况，最后通过ARCH-LM方法检验序列都存在高阶ARCH效用，可以用GARCH族建模。

关于方差方程的的移动平均阶数和自回归的阶数选取，金融实证中最常用的是GARCH(1,1)和GARCH(2,1)两种。厦门大学的郑振龙教授和中山大学陈浪南教授均针对股指在文献中指出，提高阶数并不能有效改进GARCH(1,1)模型，为此并便于比较，对GARCH族进行模拟时都不考虑高阶移动平均自回归的情况。收益序列不存在序列相关性，均值方程因此不需要引入自相关的描述部分。参数估计和检验结果，如下列各表所示：

沪深300

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | GARCH(1,1) | GARCH-M(1,1) | TGARCH(1,1) | EGARCH(1,1) |
|  | 2.70764613935e-06 | 2.72342218591e-06 | 2.93159234383e-06 | -0.169288254561 |
|  | 0.0499732865677 | 0.050102072423 | 0.0468734755085 | 0.118322761853 |
|  | 0.943362995624 | 0.94319070391 | 0.941946405293 | 0.990013443357 |
|  |  | 0.0230557060305 | 0.00719222489691 | -0.00671581706522 |
| AIC | -5.217651 | -5.216580 | -5.216765 | -5.218405 |
| SC | -5.205433 | -5.201308 | -5.201493 | -5.203132 |

香港HSI

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | GARCH(1,1) | GARCH-M(1,1) | TGARCH(1,1) | EGARCH(1,1) |
|  | 1.51775958246e-06 | 1.53165579217e-06 | 2.19011507649e-06 | -0.257660525063 |
|  | 0.0813495151006 | 0.0814013290186 | 0.0405673095105 | 0.163337505297 |
|  | 0.914255503657 | 0.914117979461 | 0.911388283158 | 0.984699300609 |
|  |  | 0.0210379208107 | 0.0758802288588 | -0.0581998321333 |
| AIC | -5.751072 | -5.750047 | -5.760073 | -5.762297 |
| SC | -5.739029 | -5.734994 | -5.745019 | -5.747244 |

韩国KSII

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | GARCH(1,1) | GARCH-M(1,1) | TGARCH(1,1) | EGARCH(1,1) |
|  | 3.13434764387e-06 | 3.1397079771e-06 | 5.97480299218e-06 | -0.445984446966 |
|  | 0.087095681386 | 0.0872084839785 | -0.00137841865989 | 0.169843668258 |
|  | 0.898531996252 | 0.898376596235 | 0.879564483892 | 0.963957385101 |
|  |  | 0.028484370796 | 0.169412307047 | - 0.122176673269 |
| AIC | -5.870359 | -5.869345 | -5.898417 | -5.897994 |
| SC | -5.858396 | -5.854391 | -5.883463 | -5.883041 |

日本N225

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | GARCH(1,1) | GARCH-M(1,1) | TGARCH(1,1) | EGARCH(1,1) |
|  | 4.11774110056e-06 | 4.11789757624e-06 | 5.48636190488e-06 | -0.464118449039 |
|  | 0.129752050961 | 0.129704045 | 0.0193629653236 | 0.2018815395 |
|  | 0.856939413353 | 0.856989376973 | 0.866754074576 | 0.964648831165 |
|  |  | 0.00062239264007 | 0.170022423805 | - 0.122837259511 |
| AIC | -5.810994 | -5.809894 | -5.839857 | -5.844401 |
| SC | -5.798886 | -5.794760 | -5.824722 | -5.829266 |

 对比GARCH模型族建模后的数据：

1. 各模型的解释效果：各AIC和 SC差别不大，都比较小。估计的参数都在水平0.05以下显著，通过显著性检验，结果都不存在序列相关性和ARCH效应。GARCH族模型较为适合。
2. 波动性的记忆效果:GARCH项系数都较大，说明股指波动都有长期记忆性。在GARCH模型中，ARCH项与GARCH项的系数之和都接近于1。表明股指波动性的记忆具有持续性，说明股市对外部冲击的反应影响时间比较长，一旦出现大的波动，短期内很难消除。其中沪深300的比其它三个股指明显偏大，ARCH项与GARCH项的系数之和最接近于1，说明中国沪深股市的长期记忆性最为明显，持续性也最显著，外部冲击对中国股市的影响最为持久。
3. 杠杆效应：在TGARCH模型中，日本和韩国的杠杆效应项的系数均显著大于零，说明这两国的股指的波动具有杠杆效应，“利空消息”能比等量的“利好消息”带来更大的波动。其中中国沪深300的系数最小，接近于零，说明中国股市的杠杆效应比较低，并不明显。这一点由EGARCH模型中的非对称项的系数的所印证。根据EGARCH模型的结果，绘制相应的信息冲击曲线。

信息冲击曲线

 

沪深300 香港HSI

 

韩国KSII 日本N225

 从信息冲击曲线可以看出，四个股指收益率在信息冲击小于零时，也就是“利空消息”负冲击时，比较陡峭，而正冲击时比较平缓。这说明负冲击使得波动性的变化更大一些。沪深300指数为代表的内地股市杠杆效应是最小的，可能与国内公司的资本结构中负债比例、股票市场的交易限制（涨跌停、禁止卖空）和开放程度、投资者成熟度有关。