

姓 名	梁依婷
学 号	202012190106

成 绩	
评阅人	

中南财经政法大学

研究生课程论文



课程名称 : 应用时间序列分析

论文题目 : 基于 Realized Garch 模型对贵金属产品的波动性分析
——以黄金 T+D 为例

专业班级 : 2020 级应用统计 2 班

学 院 : 统计与数学学院

完成时间 : 2021. 01. 07

基于 Realized Garch 模型对贵金属产品的波动性分析

——以黄金 T+D 为例

摘要：贵金属交易是一种在经济大环境不理想情况下的通用避险和资产保值手段。随着金融创新的发展以及金融危机的频发，贵金属产品的波动特征和潜在风险是研究金融交易市场的关键所在。一直以来，黄金的货币属性较强，避险、保值和抗通胀特性突出，因此，黄金的交易可看作贵金属市场变动的“晴雨表”，其价格波动可实时反映贵金属市场变化以及宏观股市环境的动态。本文选取黄金 T+D 代表贵金属现货市场，收集黄金 T+D 的 5 分钟高频交易数据，利用 5 分钟高频收盘价数据计算已实现测度 RV ，即波动估计量。首先对黄金波动特征进行研究分析，基于已实现测度 RV 和收益率数据进行描述性分析和基本检验，发现黄金的对数收益率序列均具有尖峰厚尾和波动聚集特征。然后基于不同分布假设，构建一元 Realized Garch 模型，对不同模型参数估计结果进行样本内极大似然函数值比较和样本外 Kupiec 失败率检验法 VaR 预测效果比较，选择最优模型。研究发现对黄金 T+D 对数收益率序列，在广义双曲线分布假设下基于 RV 作为已实现测度方式的 Realised GARCH(1,1) 拟合效果最好。根据最优模型结果，发现黄金收益率存在杠杆效应，负向收益率冲击对波动率的影响要比正向收益率对波动率的影响更大。

关键词：黄金 T+D；Realized Garch 模型；波动性；VaR 预测

目 录

1 引言	1
2 国内外文献综述	1
2.1 贵金属市场波动性研究综述	1
2.2 已实现波动率研究综述	2
2.3 Realized GARCH 类模型文献综述	2
2.4 文献评述	2
3 已实现波动率及 Realized Garch 类模型基础介绍	3
3.1 已实现波动率	3
3.2 Realized Garch 模型	3
4 基于 Realized Garch 模型对贵金属市场波动性分析	4
4.1 贵金属市场对象选取和数据预处理	4
4.1.1 研究对象的选取	4
4.1.2 数据预处理	4
4.2 贵金属高频数据描述性分析及检验	5
4.2.1 描述性统计分析	5
4.2.2 基本检验	6
4.3 基于 Realized Garch 模型对黄金的波动率建模	8
4.3.1 不同分布下 Realized Garch 模型参数估计	8
4.3.2 Realized GARCH 模型评价	9
4.3.3 最优模型形式	9
5 结论与启示	10
参考文献	11

1 引言

贵金属市场是一个既具备商品属性，又兼备金融属性的市场，作为金融投资商品，不仅对大宗商品交易市场有影响，甚至影响全球经济的稳定性。黄金等贵金属由于其稀缺性、价值稳定性及交易便利性受到投资者喜爱，其交易活动也日趋活跃。与股票债券不同，黄金现货即期合约采用 T+0 交易方式，可随时建仓，也可随时平仓；黄金现货延期交收合约采用 T+D 交易方式，交易方式为分期付款，交易方并不被要求必须当日交割，其允许延期交割，且无限期；如若夜间国际金价发生巨大波动，夜盘交易有利于回避其现象对国内不利的影响，增加了投资机会；保证金交易，提高资金利用率^[1]。

2019 年 12 月底新冠病毒在中国被发现，即使各地政府极力防护，但最终也演变为蔓延全世界的重大事件，世界经济处于自 2008 年全球金融危机以来最不稳定的状态，不可避免的造成全球经济的“停摆”。2020 年 3 月，我国疫情接近尾声，而新冠肺炎在欧美国家大面积爆发，美股在十天内发生了四次熔断。面对通货膨胀压力，不断加剧的金融市场风险，投资者对资金的避险需求不断提高，贵金属市场尤其是黄金市场的保值增值能力，引起了投资者的关注。金融市场动荡下，贵金属市场的走势如何，贵金属产品的波动性会有何特征值得探究。

然而，随着现代金融市场的快速发展，市场波动变化相比与以往更难把握，高频数据在对市场波动等相关信息的呈现远远超过传统的低频数据。科技水平的提高使得高频数据的获取成为可能，如果能充分利用数据自身包含的信息，将高频数据与易于估计的模型相结合，就可以提高对产品波动性的估计精度。

本文以黄金代表贵金属现货市场，选取 2019 年 7 月 1 日至 2020 年 12 月 31 日黄金 T+D 高频价格数据作为研究对象，对贵金属产品的波动性进行研究。通常，金融时间序列表现为尖峰厚尾的特征，其收益率一般不服从正态分布，故本文采用 Realized GARCH 模型对残差服从正态分布，t 分布，偏 t 分布、广义双曲线分布等对黄金的收益率序列进行拟合，根据极大似然函数值和风险价值预测选出拟合效果较好的分布，并对其波动特征进行分析。

2 国内外文献综述

2.1 贵金属市场波动性研究综述

波动性意味着价格在某一时间间隔所发生的变化，与投资者的收益息息相关。之前对于贵金属市场波动性的研究，学者采取了不同的模型及研究方法。林新丹 (2014) 通过对我国黄金和白银期货价格的相关性研究结果说明高频数据更有助于对黄金、白银等贵金属价格的预测，2013 年黄金大跌对白银市场造成了巨大的冲击，影响程度达到 90% 以上^[2]。杨胜刚等 (2019) 对伦敦金属交易所的四种贵金属 (黄金、白银、铂金和钯金) 基于 time-varying AR 模型构建有效性程度指标对适应性市场假说 (AMH) 进行实证研究，发现黄金市场有效性程度指标的均值和波动幅度相对最小，而钯金市场的均值和波动幅度则相对最大^[3]。

2.2 已实现波动率研究综述

在现代金融市场中，波动率是一种可以反映金融资产价格变化的统计度量，准确度量 and 预测金融资产波动率，在金融资产定价，金融风险控制等决策中是不可或缺的一步。传统对金融市场的波动率研究大多基于日、周、月等低频交易数据，如 GARCH 类和 SV 类模型。随着信息技术的不断提高，高频数据的易获取使得人们对波动率的研究更进一步。

Merton(1980) 在研究市场预期收益时首次提出已实现波动率的概念，将一个时间段划分为多个子时间段区间，用多个子时间段的收益率平方和估计该时间段的收益率方差^[4]。已实现波动率概念提出以后，由于当时的技术限制，划分数据子区间还不能实现，无法真正应用到实践中。Andersen 和 Bollerslev 等人 (1998) 首次将高频数据应用于波动率的测度上，提出已实现波动率 (RV, Realized Volatility) 的估计计算方法^[5]，与传统的根据历史日间交易数据估计波动率不同，这种方法直接利用“日内收益平方和”估计波动率，证明了当抽样频率非常大时，已实现波动是积分波动的一致估计量。

2.3 Realized GARCH 类模型文献综述

已实现波动率是把高频数据应用到波动率领域，使得金融资产的波动率能够直观表示。Hansen 等人 (2012) 提出了 Realized GARCH 模型，将已实现波动率与 GARCH(广义自回归条件异方差模型) 结合，通过一个测量方程连接条件方差和已实现波动率，并在方程这种加入杠杆函数体现非对称效应，把已实现波动率作为条件方差的估计，实现收益率，波动率和已实现测度的联合建模，和传统 GARCH 类模型相比，Realized GARCH 模型提高了拟合能力和波动性的预测能力^[6]。后来很多学者将 Realized GARCH 模型进行了各种各样的延伸，并广泛应用于一些金融实证研究中。下面主要对一元 Realized GARCH 模型的发展、应用等进行叙述。

由于金融时间序列往往不具备正态分布的特征，Watanabe(2012) 将 Hansen(2012) 提出的残差服从正态分布的 Realized GARCH 模型扩展为残差服从 t 分布，偏 t 分布的 Realized GARCH 模型，并对基于这三种分布下的 Realized GARCH 模型预测 VaR 和 ES 进行对比研究^[7]。Huang 等人 (2016) 对 S&P500 指数期权进行实证研究，表明基于 Realized GARCH 模型算出的定价误差小于传统模型^[8]。

2.4 文献评述

从贵金属市场研究文献看，国内学者对贵金属的研究主要集中在黄金、白银、铂金上，更为常见的是集中在对黄金市场的相关分析上。因此，本文选取黄金代表贵金属市场进行波动性研究。

已实现波动测度是基于高频数据计算的，阅读文献可知 Realized GARCH 类模型有两个基本条件：高频数据的获取以及已实现波动率的计算及最优选择。高频数据能够反映更多信息，相比于低频日数据，基于高频数据对波动率等的拟合和预测性能更佳。对于一元 Realized GARCH 模型，针对金融时间序列尖峰厚尾的特性，将 Realized GARCH 模型残差分布从正态分布扩展到 t 分布，偏 t 分布、广义双曲线分布等，通过比较不同分布下的预测效果选择最优模型。

3 已实现波动率及 Realized Garch 类模型基础介绍

3.1 已实现波动率

已实现波动率是基于高频数据度量的金融资产波动率。假定不考虑由买卖价差、闭市效应、非同步交易等因素，且没有发生跳跃，按照 Andersen 和 Bollerslev(1998) 对已实现波动率的计量，已实现的波动率 RV 是计算期第 t 日观测到的收益率的平方和，每日交易被分割为 M 段，当 M 趋向于无穷大时

$$RV_t = \sum_{i=1}^M r_{t,i}^2 \quad (1)$$

$r_{t,i}$ 为金融资产或者资产组合在第 t 交易日内的第 i 个间隔的日内对数收益率， $\ln P_{t,i}$ 表示金融资产或者资产组合在第 t 个交易日内的第 i 个间隔的日内对数价格， T 表示计数天数， M 表示每天的时间间隔为 M ，与采样频率有关^[5]。

3.2 Realized Garch 模型

普通的 GARCH 类模型一般是通过每日收益获得相应的波动率结果，但是波动率是通过滞后多阶的条件方差计算得到的，所以当收益情况波动剧烈时，GARCH 类模型较难把所有信息有效利用。Hansen 等 (2012) 提出 Realized Garch 模型，该模型用波动率的已实现测度代替对未来波动率进行建模^[6]。

根据 Hansen 等 (2012) 的研究结果,Realized Garch 模型在一阶下的估计效果已经相当不错，所以本文主要采用 Realized GARCH(1,1) 模型进行实证研究，该模型具体形式为

$$\begin{cases} r_t = \sqrt{h_t} z_t, z_t \sim N(0, 1) \\ \ln h_t = \omega + \beta \ln h_{t-1} + \alpha \ln x_{t-1} \\ \ln x_t = \xi + \varphi \ln h_t + \tau(z_t) + u_t, u_t \sim iid(0, \sigma_u^2) \end{cases} \quad (2)$$

模型的三个等式分别为收益方程、GARCH 方程和测量方程。 r_t 是第 t 日的日收益率， $h_t = \text{var}(r_t | F_{t-1})$ 表示收益率的条件方差， x_t 表示第 t 日波动率的已实现测度，本文使用 RV 作为已实现波动率的估计。 z_t 是收益率的标准随机误差项，残差项 $u_t \sim iid(0, \sigma_u^2)$ ，残差项可服从正态分布、t 分布、偏 t 分布、广义双曲线分布等。 z_t 与 u_t 相互独立。测量方程将 h_t 和 x_t 联系起来。在 Realized GARCH 模型中 h_t 是 h_{t-1} 和 x_t 的函数。

考虑到收益率冲击对波动率具有一定的非对称影响，本文在测量方程中引入了杠杆函数 $\tau(z_t)$ ，Hansen 等 (2012) 人将 $\tau(z_t)$ 设定为以下二次形式

$$\tau(z_t) = \tau_1 z_t + \tau_2 (z_t^2 - 1) \quad (3)$$

杠杆函数描述了 Realized GARCH 模型的信息冲击曲线，表明过去时刻的收益率在大小和方向上同时都影响着波动率，即正的价格扰动对波动的影响和负的价格扰动对波动率产生的影响在大小和方向上都不一致的。

4 基于 Realized Garch 模型对贵金属市场波动性分析

4.1 贵金属市场对象选取和数据预处理

4.1.1 研究对象的选取

贵金属主要包含金、银和铂族金属等金属元素，我国贵金属的现货交易主要在上海黄金交易所进行。通过计算上海黄金交易所中近 2 年产品月成交量占比可发现，黄金延期 Au(T+D) 的月成交量占比基本在 95% 以上，且其上市时间也相对较久，因此本文将选择黄金延期 Au(T+D) 合约作为研究对象，以代表贵金属现货市场进行研究。

现货交易分日市交易和夜市交易，一般情况（即非国家法定节假日）下，日间交易时间为星期一到星期五的 09:00-15:30，夜市交易时间为星期一到星期五的 20:00-02:30。所以不能按传统的交易日来看，即当天发生的交易为一个交易日来进行研究。对本文研究的对象 Au(T+D) 来说，当交易日非国家法定节假日时，前日 20:00 对应交易日的开盘价，当日 15:30 对应交易日收盘价；当交易日为国家法定节假日时，后日 9:00 为开盘时间，15:30 为收盘时间。

4.1.2 数据预处理

本文选择 5 分钟的频率作为高频数据的抽样频率，选取 Au(T+D) 的 5 分钟交易数据，样本区间为 2019 年 7 月 1 日晚八点至 2020 年 12 月 31 日下午三点半，数据源自东方财富旗下的 Choice 金融终端数据库。

由于黄金含有夜市交易和日市交易，因此将数据分为两类处理。国家法定节假日后的第一天交易数据从当日 9:00 至 15:30 的交易形成，每天有 80 个时间间隔，其他交易日交易时间为前一日 20:00 至当日 02:30, 9:00 至 15:30，每天有 158 个时间间隔。由于开盘价是集中竞价¹的结果，所以剔除 9:00, 20:00 的数据。将数据分为两组，第一组为 2019 年 7 月 1 日至 2020 年 6 月 30 日，作为样本内数据用于拟合模型，第二组为 2020 年 7 月 1 日至 2020 年 12 月 31 日，作为样本外数据，用于对比检验模型效果。对于交易数据的处理，剔除当天停牌或者价格几乎不变化的交易日。经过数据处理后，Au(T+D) 有 367 个交易日，共有 56160 个 5 分钟的高频数据。

考虑到隔夜价格的影响，收益率计算方式将依据“当日收盘价-前一日收盘价”。设 t 为任意一个特定的交易日， $p_{t,j}$ 表示第 t 日的第 j 个区间的收盘价，在第 t 个交易日中，第 j 个区间的对数收益率为

$$r_{t,j} = \ln p_{t,j} - \ln p_{t,j-1} \quad (4)$$

¹集中竞价有利于提高交易效率，降低交易成本，减少交易纠纷。

4.2 贵金属高频数据描述性分析及检验

4.2.1 描述性统计分析

本节首先根据式 (4) 计算出 Au(T+D)2019 年 7 月 1 日至 2020 年 12 月 31 日的每日收益率，共 367 个收益率数据。然后根据样本区间每个交易日的以 5min 为时间间隔的收盘价分别计算出 Au(T+D) 的高频日内收益率，进而计算相关已实现测度估计量，本文采用已实现测度估计量 RV。考虑到计算出来的收益率过小，为了方便研究，将收益率结果扩大 100 倍。

4.2.1.1 序列图 Au(T+D) 收盘价和收益率时序图如图1所示，从 Au(T+D) 的日收盘价时序图中可以看出，Au(T+D) 价格整体呈上升趋势，但期间也有较大的波动。在收益率时序图中，收益率围绕着零上下波动，波动的幅度与收盘价的涨幅情况基本一致，2020 年 3-4 月和 8-9 月收益率波动幅度较大，具有明显的“波动性聚集”特征。从 2020 年 3 月开始，国外疫情逐渐严重、风险资产也愈跌愈烈、市场整体消极，黄金的避险属性由此凸显，出现价格上涨和收益率大幅度波动的情况。2020 年 8 月，各国开始陆陆续续公布新冠疫苗的进展，市场有所回暖，但也在一定程度上造成黄金的下跌，其对应的收益率也波动剧烈。

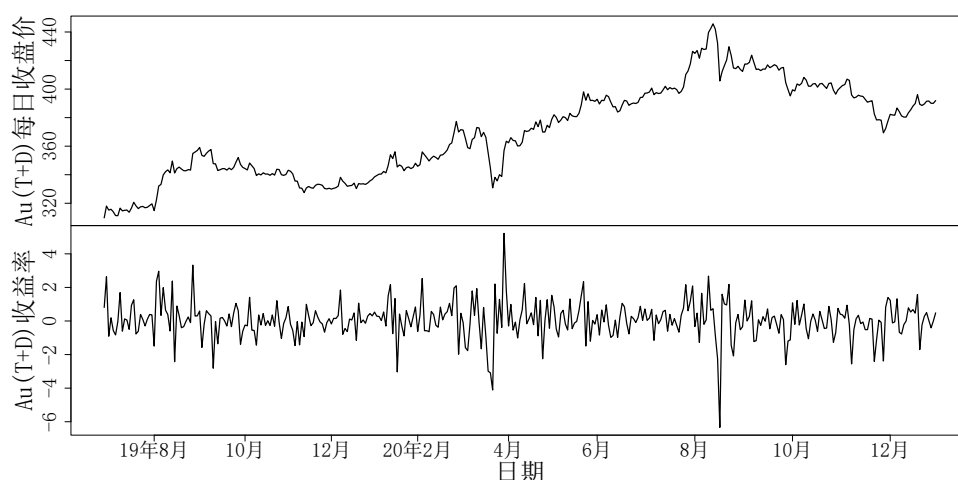


图 1 Au(T+D) 收盘价和收益率序列时序图

4.2.1.2 基本统计特征 表1列出了 Au(T+D) 对数收益率序列 r_t 和已实现测度 RV 序列的统计特征。可以明显看出， r_t 和 RV 序列的峰度分别为 4.853 和 33.296，均大于 3，可认为它们都具有高峰厚尾的特征，Au 收益率序列 r_t 偏度小于 0，说明其左偏，波动率估计量 RV 偏度均大于 0，具有右偏特征，且峰度较大。两个序列的 J-B 统计量²均较大，表现了显著不服从于正态分布的特点。

表 1 Au(T+D) 收益率和已实现测度统计特征

	均值	最小值	最大值	标准差	偏度	峰度	JB 统计量
Au_rt	0.0654	-6.337	5.219	1.096	-0.4433	4.853	378.7

²J-B 统计量又名 Jarque-Bera 统计量，可以用来检验观测对象是否服从正态分布。

表 1 Au(T+D) 收益率和已实现测度统计特征 (续)

	均值	最小值	最大值	标准差	偏度	峰度	JB 统计量
Au_RV	0.7717	0.047	10.819	1.163	5.1468	33.296	18789.3

4.2.2 基本检验

4.2.2.1 平稳性检验 为保证实证结果的可靠性，在进行实证分析之前首先要对收益率序列进行平稳性检验。本文采用 ADF 检验法，检验结果如表2所示。

表 2 Au(T+D) 收益率序列平稳性检验

	检验统计量	P 值
ADF 检验	-8.458	0.0001

通过表2的平稳性检验结果可以看到，Au(T+D) 收益率序列 r_t 的 ADF 检验的 p 值远小于显著性水平 0.01，拒绝原假设，因此可以认为 Au(T+D) 收益率 r_t 序列是平稳的。

4.2.2.2 白噪声检验 对 Au(T+D) 的收益率序列进行白噪声检验，这里分别运用 (偏) 自相关系数图和 Ljung-Box Q 统计量对 Au(T+D) 的收益率序列进行分析。

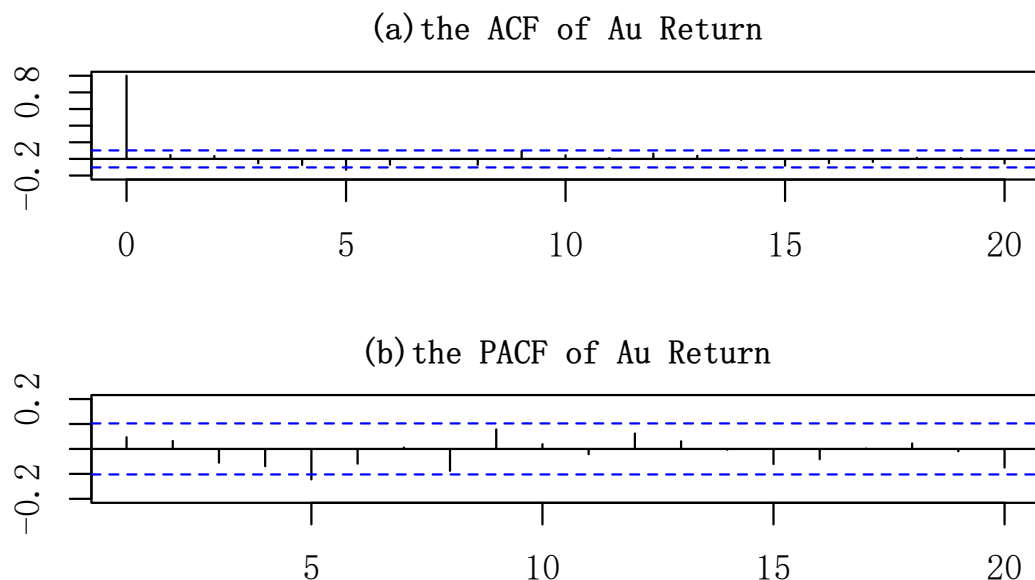


图 2 收益率自相关图和偏自相关图

图2显示了 Au(T+D) 的收益率序列 r_t 的自相关图和偏自相关图，从图中可以看出，滞后阶数为 20 阶以内下，自相关系数均未超出 95% 置信区间，表明序列不具有自相关性；偏自相关系数也均未超出 95% 置信区间，表明序列不具有偏自相关性。

表 3 Au(T+D) 收益率序列纯随机性检验

	LB 统计量	延迟阶数	P 值
纯随机性检验	14.62	10	0.1464

Ljung-Box 检验是用于检验给定时间序列是否存在滞后相关性的一种检验方法，原假设为序列为白噪声序列。通过表3的平稳性检验结果可以看到，Au(T+D) 收益率序列 Ljung-Box Q 统计量的值为 14.6229，对应 p 值为 0.1464，大于显著性水平 0.05，故 Au(T+D) 收益率序列为白噪声序列，不具有相关性。

综合之前的平稳性检验，可以判断对数收益率序列 r_t 是一个平稳白噪声序列，不用构建 ARMA 模型。

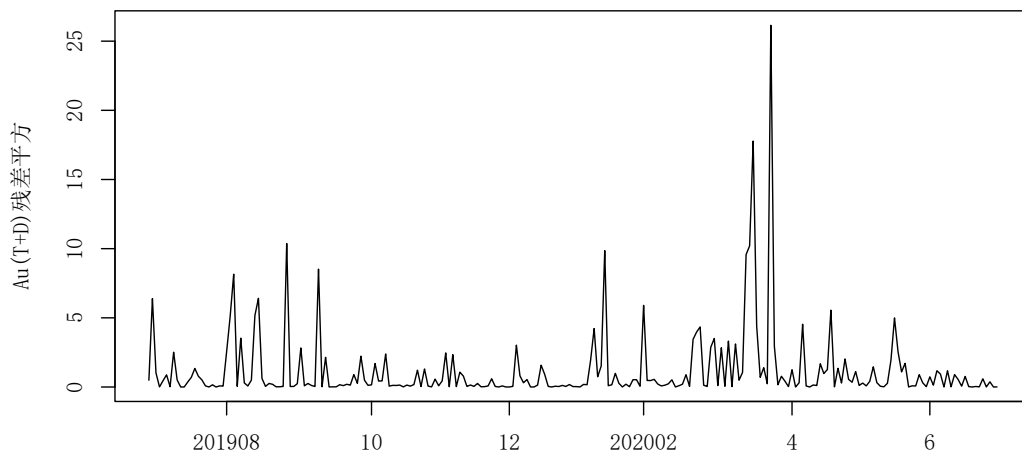


图 3 收益率残差平方时序图

4.2.2.3 残差平方序列异方差性检验 经标准化后得到的残差平方时序图如图3所示，从图中可以看出，对于 Au(T+D) 收益率去均值化后的残差依然有较显著的“波动性聚集”特征，这说明收益率序列可能存在异方差现象，需要进行异方差性检验。本文运用 ARCH 检验方法中的 LM 检验。

表 4 Au(T+D) 残差平方序列 ARCH 检验

	LM 检验统计量	自由度	P 值
ARCH LM-检验	28.2	5	0

ARCH 效应的存在会导致估计的有效性降低，其原假设为：残差序列中直到 p 阶都不存在 ARCH 效应。对标准化后的残差平方进行 ARCH 效应检验，检验结果如表4所示。Au(T+D) 的 LM 统计量的值为 28.2，对应的 P 值为 0.0000，在 0.05 的显著性水平上拒绝残差独立的原假设，表明标准化后的残差平方存在 ARCH 效应。因此需对 Au(T+D) 收益率的残差拟合 GARCH 模型，以消除 ARCH 效应。

4.3 基于 Realized Garch 模型对黄金的波动率建模

4.3.1 不同分布下 Realized Garch 模型参数估计

前文的分析指出 Au(T+D) 收益率序列具有波动特点。相较于普通的 GARCH 类模型而言, 基于高频数据构建的 Realized Garch 模型更能真实的反映序列的波动性, 且低阶 GARCH 类模型的拟合效果较好, 且在精度方面不存在太大的差距, 结果可靠。因此本文选择基于 5 分钟的高频收益率对 Au(T+D) 进行建模, 选择已实现方差 RV 作为已实现测度, 将 Realized GARCH 的阶数定为 (1,1)。

在分布类型的选取方面, 由前文的分析可知, Au(T+D) 收益率序列和已实现测度 RV 表现为偏态和尖峰厚尾的特征, 表现为非正态分布。为了准确地描述序列的厚尾性, 分布类型依次选取正态分布 (N)、对肥尾特点刻画比较灵活的学生 t 分布 (T)、偏 t 分布 (ST) 和广义双曲线分布 (GH)。选用 R 软件的 rugarch 包进行 Realized GARCH 模型的构建。

表 5 Au(T+D) 不同模型参数估计结果

模型形式	ω	α	β	τ_1	τ_2	δ	λ	ξ	π
N	0.373	0.435	0.339	-0.221	0.225	1.11	0.563	-0.822	0.819
	0.000	0.000	0.002	0.000	0.000	0.00	0.000	0.000	-
T	0.539	0.503	0.304	-0.244	0.261	1.02	0.562	-0.941	0.815
	0.007	0.000	0.006	0.000	0.000	0.00	0.000	0.000	-
ST	0.528	0.505	0.302	-0.242	0.255	1.01	0.562	-0.926	0.814
	0.008	0.000	0.006	0.000	0.000	0.00	0.000	0.000	-
GH	0.422	0.484	0.319	-0.227	0.226	1.03	0.562	-0.831	0.82
	0.001	0.000	0.005	0.000	0.000	0.00	0.000	0.000	-

注: 每个参数估计结果的第一行为参数值大小, 第二行为每个估计结果的 p 值。

表5是基于不同分布假设下应用 RV 已实现测度构建的 Au(T+D) 的 Realized GARCH 模型参数估计结果。从表中可以看出, 不同类型的 Realized GARCH 模型的所有系数在 5% 的显著性下都是显著的。不同分布模型的估计结果中 α 、 β 均满足 $\alpha > 0$ 、 $\beta > 0$ 、 $\alpha + \beta > 0$ 的条件, 说明模型比较稳定; 系数 δ 在 1 附近, 说明测量方程是恰当的; 表中 π 表示持续系数³, 持续系数稳定在 0.8 附近, 说明 Au(T+D) 的波动率存在较强的可持续性。通过对比各模型的参数可以看出, 基于不同分布的模型估计结果中的代表波动率长期记忆性参数 β 值均约为 0.3, 相互之间的差距不明显, 说明影响当期波动率的不仅有前一期的波动率, 还有前一期已实现测度的因素; 测量方程中的杠杆函数 $\tau(z_t)$ 用于描述金融资产收益率不同方向的扰动对波动的非对称影响, 从表中 $\tau_1 < 0$, $\tau_2 > 0$, 的结果可以看出, 模型负向扰动对波动的影响比正向扰动的影响大, 且这种非对称的影响较为明显。

³ $\pi = \beta + \alpha * \delta$, 用于直观说明波动率的长期记忆性, 该系数一般而言应接近于 1

4.3.2 Realized GARCH 模型评价

4.3.2.1 样本内拟合优度评价——似然函数值 判断模型拟合效果的依据为对数似然函数值的大小，其值越大，说明拟合效果越好，模型越有效。从表 6 中可以看到，在基于已实现测度 RV 构建的 Realised GARCH 模型的四种分布中，对数似然函数值从正态分布、t 分布、偏 t 分布到广义双曲线分布逐渐增大，广义双曲线分布假设下构建的模型极大似然函数值最大，可以认为在广义双曲线分布假设下基于 RV 作为已实现测度方式的 Realised GARCH 模型拟合度最好。

表 6 Au(T+D) 不同分布下的模型极大似然函数值

	正态分布	t 分布	偏 t 分布	广义双曲线分布
对数似然函数值	-551	-539	-539	-536

4.3.2.2 样本外预测效果评价——风险价值 为了使分析更为全面，本节选取 Kupiec 失败率检验法在 95% 置信水平下对各模型的预测效果进行比较，采用的预测方法为固定窗口一步向前法，估计对象为样本外 VaR 值⁴。预测区间为 2020 年 7 月 1 日至 2020 年 12 月 31 日，共 126 个交易日。LR 越小，p 值越大，表明模型越精确，可信度越高。

表 7 不同 Realized Garch 模型样本外 VaR 一步预测

VaR 显著性水平	模型形式	LR 统计量	失败率	失败天数	P 值
$\alpha=0.01$	N	1.08	0.071	9	0.298
	T	1.96	0.079	10	0.162
	ST	1.96	0.079	10	0.162
	GH	1.08	0.071	9	0.298

表 7 展示了基于不同分布构建的 Au(T+D) 的 Realized Garch 模型样本外 VaR 一步预测结果。从表中可以看出 p 值较大，不能拒绝原假设，这说明构建的这些模型对 VaR 预测结果都是有效的。置信水平为 95%(显著性为 0.05) 时，基于广义双曲线分布假设构建的模型效果也是相对更好的。

综合样本内对拟合优度评价和样本外预测效果评价可以得出以下结论：对于 Au(T+D) 收益率序列，在广义双曲线分布假设下基于 RV 作为已实现测度方式的 Realised GARCH 模型拟合度最优，风险预测效果最好。

4.3.3 最优模型形式

基于以上的分析，对于 Au(T+D) 收益率序列选择在广义双曲线分布假设下基于 RV 作为已实现测度方式的 Realised GARCH(1,1) 模型，代入估计的参数可

⁴风险价值 (Value at Risk, VaR) 是用来测量金融资产在市场价格波动的情况下可能的最大损失。

得最优模型的形式如下：

$$\begin{cases} r_t = \sqrt{h_t} z_t \\ \log h_t = 0.422 + 0.319 \log h_{t-1} + 0.484 \log x_{t-1} \\ \log x_1 = -0.831 + 1.03 \log h_t + \tau(z_t) + \mu_t \\ \tau(z_t) = -0.227 z_t + 0.226 (z_t^2 - 1) \end{cases} \quad (5)$$

r_t 为 Au(T+D) 的日对数收益率, z_t 服从广义双曲线分布, h_t 为隐含波动率, x_t 表示波动率的已实现度量, 这里使用的是 RV 作为已实现波动率的估计。

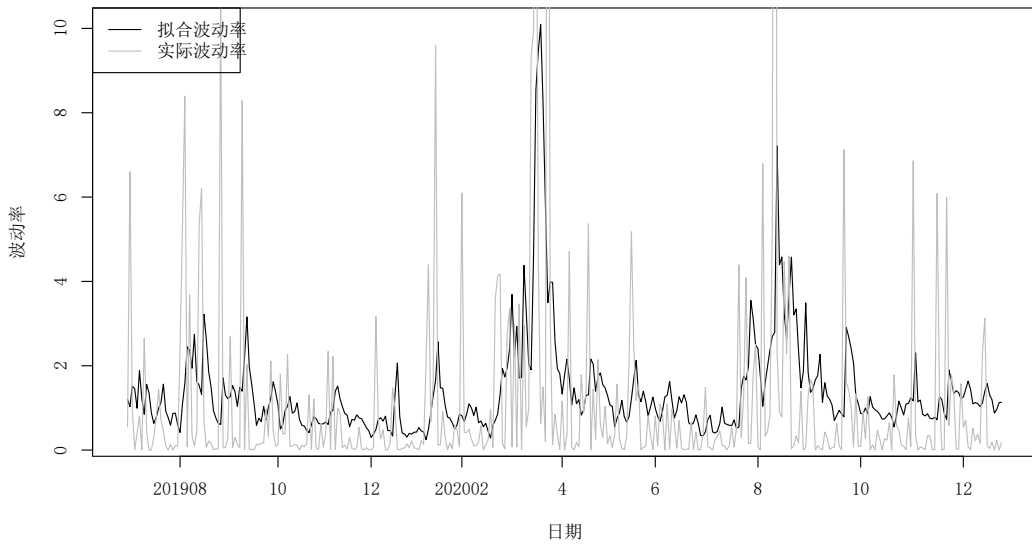


图 4 Au(T+D) 波动率拟合图

利用最优模型对 Au(T+D) 波动率进行拟合, 结果如图4所示。通过对不同时刻波动率的表现特征分析发现, 自 2019 年 8 月到 9 月初, Au(T+D) 出现较小幅度的波动, 在 2019 年 7 月底-8 月初英国强硬脱欧、贸易战愈演愈烈, 边缘战争不断爆发等事件的发生点燃了黄金行情。2019 年 8 月以来许多货币对美元出现了贬值, 美债收益率倒挂严重, 在美国实际利率下行趋势开启的背景下, 投资者对经济衰退的担忧加剧, 由于黄金具有套期保值的特性, 黄金价格逐渐上涨。2020 年 3 月开始 Au(T+D) 出现了剧烈波动, 结合当时现实背景可知, 当时新冠疫情在世界大面积爆发, 美国、英国等国家逐渐出现疫情失控的局面, 市场持悲观态度, 市场风险有所增加, 这也引起了黄金价格的大幅度上涨, 引起了剧烈波动。2020 年 8-9 月, 各国开始陆陆续续公布新冠疫苗的进展, 市场有所回暖, 但也在一定程度上造成黄金的下跌, 其对应的收益率也波动剧烈。

5 结论与启示

本文选取了在贵金属市场中具有代表性的黄金作为研究对象, 以 Au(T+D) 的 5 分钟高频收益率序列, 结合已实现测度 RV 构建一元 Realized GARCH 模型, 对黄金波动特征研究分析, 可获得如下结论:

1. Au(T+D) 的收益率序列呈尖峰厚尾的现象, 不具有正态分布的特征, 且存在显著的条件异方差效应, 即在贵金属现货市场中黄金的波动表现出显著的时变性和集聚性。
2. 利用 5 分钟高频收益率计算 Au(T+D) 的已实现方差 RV, 基于不同分布构建的一元 Realized GARCH 模型后, 发现 Au(T+D) 基于 RV 和广义双曲线分布的 Realized GARCH 拟合效果最好, 说明已实现方差 RV 能捕捉到连续的价格突变, 减少高频交易数据的微观噪音。
3. 根据测量方程中的杠杆函数 $\tau(z_t)$ 所对应的 $\tau_1 < 0$, $\tau_2 > 0$, 得出收益率序列存在明显的非对称性, 即杠杆效应, 并且负向收益率冲击对波动率的影响要比正向收益率对波动率的影响更大。
4. 通过分析最优模型刻画的波动率, 发现黄金在市场动荡以及新冠疫情影响下波动剧烈, 在市场动荡, 趋于不利背景下, 贵金属的保值避险优势尽显。

参考文献

- [1] 郭凡礼. 国内贵金属交易市场渐入佳境[J]. 中国银行业, 2015(05):111-113.
- [2] 林新丹. 2013 年黄金大跌对白银市场的影响力探讨[J]. 特区经济, 2014(06): 98-99.
- [3] 杨胜刚, 彭啸帆, 邹子昂, 等. 适应性市场假说的实证研究——基于国际现货贵金属市场[J]. 湘潭大学学报 (哲学社会科学版), 2019, 43(03):82-88.
- [4] Merton Robert C. On estimating the expected return on the market: An exploratory investigation[J]. North-Holland, 1980, 8(4).
- [5] Andersen T, Bollerslev T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts[J]. International Economic Review, 1998, 39.
- [6] Hansen P R, Huang Z, Shek H H. Realized garch: a joint model for returns and realized measures of volatility[J]. Journal of Applied Econometrics, 2012, 27(6).
- [7] Watanabe T. Quantile Forecasts Of Financial Returns Using Realized Garch Models[J]. The Japanese Economic Review, 2012, 63(1):68-80.
- [8] Huang Z, Wang T, Hansen P. Option pricing with the realized garch model: An analytical approximation approach: Option pricing with the realized garch model[J]. Journal of Futures Markets, 2016, 37.

附录（R 代码）：

```
###加载包
library(kableExtra)
library(knitr)
library(fBasics)
library(tseries)
library(xts)
library(readxl)
library(urca)
library(FinTS)
library(parallel)
library(rugarch)
library(fUnitRoots)
library(fGarch)
library(highfrequency)

#####构建已实现测度函数#####
Rt=function(x) {
  rt=rep(0,length(x)) #x 为收盘价向量
  for (i in 2:length(x)) {
    rt[i]=(log(x[i])-log(x[i-1]))*100
  }
  rt
} # 计算日内收益率函数,将收益率扩大了 100 倍

RV=function(x,y,z) {
  rv=rep(0,length(x)) #x 为交易日向量
  m=length(y)/length(x) #y 为收盘价向量, m 为时间间隔
  for (i in 1:length(x)) {
    for (j in 1:m){
      rv[i]=rv[i]+z[m*(i-1)+j]^2 #z 为收益率向量
    }
  }
  rv
} #计算已实现波动 RV

#####计算收益率#####
path="./data/data_raw"
fileNames=dir(path)
Lst1<-list() # 日内收益率数据
for(i in fileNames) {
  Lst1[[i]]<-as.data.frame(read.csv(paste(path,i,sep = "/"))) #
  读入数据
}
```

```

Lst1[[i]]$DATE=as.character(Lst1[[i]]$DATE)
Lst1[[i]]$TIME=as.character(Lst1[[i]]$TIME) #转换数据
Lst1[[i]]$rt=Rt(Lst1[[i]]$Close) #计算日内收益率
Lst1[[i]]=Lst1[[i]][which(!(Lst1[[i]]$TIME %in% c("9:05","20:
05"))),] #剔除异常数据
}

names(Lst1) # "Au1.csv": 普通时间; "Au2.csv": 节假日后第一天
#####各种已实现测度#####
#### 计算交易日和已实现测度

Lst3=list() #存放各种向量和已实现测度表
for (i in (1:2)) {
  Lst3[[i]]=Lst1[[i]][which(Lst1[[i]]$TIME == "15:30"), "DATE"]
  #计算交易日
  Lst3[[i+2]]=RV(Lst3[[i]],Lst1[[i]]$Close,Lst1[[i]]$rt) #已实现
波动RV
  Lst3[[i+4]]=as.data.frame(cbind(Lst3[[i]],Lst3[[i+2]]),string
sAsFactors = F) ##各种已实现测度
  colnames(Lst3[[i+4]])=c("date","RV")
}
names(Lst3)[1:2]=c("Au1date","Au2date") #交易日
names(Lst3)[3:4]=c("Au1_rv","Au2_rv") #RV
names(Lst3)[5:6]=c("Au1real","Au2real") #已实现测度表

Au_dayrt=read_excel('./data/day_rt.xlsx',sheet=1)#读入日收益率
Au_dayrt$date=as.Date(Au_dayrt$date)

Real_Au=rbind(Lst3[[5]],Lst3[[6]]) # Au 所有交易日已实现测度
Lst4=list(Real_Au,Au_dayrt) #存放合并后的已实现测度
names(Lst4)[1]=c("Real_Au.csv")
Lst4[[1]]$date=lubridate::ymd(Lst4[[1]]$date) #转为日期形式
Lst4[[1]]=dplyr::arrange(Lst4[[1]], date) # 按日期排序
Lst4[[1]]$RV=as.numeric(Lst4[[1]]$RV)
Lst4[[1]]=merge(Lst4[[1]],Lst4[[2]],sort = FALSE,all = TRUE) #加入
日收益率和日收盘价
write.csv(Lst4[[1]],paste('./data/',names(Lst4)[1],sep = ""),row.
names = F) #保存

#读入数据
AU=read.csv("./data/Real_Au.csv",header=TRUE,sep=',',stringsAsFac
tors=FALSE)
AU$date=as.Date(AU$date)

```

```

par(mfrow=c(2,1))
par(mai=c(0,2,1,1))
plot(AU$close,type='l',xlab='日期',ylab='Au(T+D) 每日收盘价',
      ,xaxt="n",cex.lab=2.5,cex.axis=2,lwd=2)
par(mai=c(1,2,0,1))
plot(AU$rt,type='l',xlab='日期',ylab='Au(T+D) 收益率',
      ,xaxt="n",las=0,cex.lab=2.5,cex.axis=2,lwd=2)
axis(1,labels=c("19年8月","10月","12月","20年2月","4月","6月","8月",
"10月","12月"),at=c(23,63,101,139,179,218,261,304,347),cex.axis=2)

```

```

Real_Au <-xts(read.zoo("./data/Real_Au.csv",header=TRUE,sep=','))

```

####Au 基本统计特征

```

Au_rt=basicStats(Real_Au$rt) #rt 基本统计特征
Au_RV=basicStats(Real_Au$RV) #RV 基本统计特征

```

```

Au_rt_JB=jarque.bera.test(Real_Au$rt)#(p<0.05 非正态) #JB 检验
Au_RV_JB=jarque.bera.test(Real_Au$RV)

```

```

Auchara<-as.data.frame(array(,dim=c(2,7)))
Auchara[1,<-
t(c(Au_rt[c("Mean","Minimum","Maximum","Stdev","Skewness","Kurtosis"),1],as.numeric(Au_rt_JB[[1]][1])))
Auchara[2,<-
t(c(Au_RV[c(7,3,4,14,15,16),1],as.numeric(Au_RV_JB[[1]][1])))
colnames(Auchara)=c('均值','最小值','最大值','标准差','偏度','峰度','JB 统计量')
rownames(Auchara)=c('Au_rt','Au_RV')

```

####Au 收益率序列平稳性检验#####

```

AuADF=adf.test(Real_Au$rt) #ADF 检验
Auadf_test=data.frame(a=AuADF$statistic[1],b=0.0001)
colnames(Auadf_test)=c("检验统计量","P 值")
rownames(Auadf_test)="ADF 检验"

```

```

par(mfrow=c(2,1))
par(mar=c(5,4,2,1))
#Au
acf(Real_Au$rt,main="",xlab="滞后期",
    ,ylab="ACF",lag.max=20,ylim=c(-0.2,1))#画自相关图
title(main="(a) the ACF of Au Return",cex.main=0.95)
pacf(Real_Au$rt,main="",xlab="滞后期",
     ,ylab="PACF",lag.max=20,ylim=c(-0.2,0.2))#画偏自相关图

```

```

title(main="(b) the PACF of Au Return",cex.main=0.95)

#####Au 收益率序列-LB 统计量#####
Au_box=Box.test(Real_Au$r[1:241],lag=10,type = "Ljung") #随机性检验
Q_test=data.frame(a=Au_box[1],df=Au_box[2],p=Au_box[3])
colnames(Q_test)=c("LB 统计量", "延迟阶数", "P 值")
rownames(Q_test)="纯随机性检验"

par(mfrow=c(1,1))
par(mar=c(2,4,1,2))
x=c("201908", "10", "12", "202002", "4", "6", "8", "10", "12")
xat=c(23,64,103,141,183,222,265,308,351)
fit1=stats::arima(Real_Au$r[1:241],order = c(0,0,0))
Au_r<-stats::residuals(fit1)
plot(Au_r^2,type="l",
      xlab="日期",ylab="Au (T+D) 残差平方",xaxt="n")
axis(1,labels=x,at=xat)

#####Au 收益率序列-ARCH 检验#####
Au_arch=ArchTest(Au_r^2,lag=5) #异方差性检验
arch_test=data.frame(stat=Au_arch[1],df=Au_arch[2],p=0.0000)
colnames(arch_test)=c("LM 检验统计量", "自由度", "P 值")
rownames(arch_test)="ARCH LM-检验"

Real_Au1 <-
xts(read.zoo("./data/Real_Au.csv",header=TRUE,sep=', '))
Real_Au=Real_Au1[1:241] #样本内参数估计

###设定模型分布形式
norm_spec.mod= ugarchspec(
  variance.model = list(model = "realGARCH", garchOrder = c(1,1))
,
  mean.model = list(armaOrder = c(0,0), include.mean = FALSE),
  distribution.model = "norm") # 正态分布
std_spec.mod= ugarchspec(
  variance.model = list(model = "realGARCH", garchOrder = c(1,1))
,
  mean.model = list(armaOrder = c(0,0), include.mean = FALSE),
  distribution.model = "std") # t 分布
sstd_spec.mod= ugarchspec(
  variance.model = list(model = "realGARCH", garchOrder = c(1,1))
,
  mean.model = list(armaOrder = c(0,0), include.mean = FALSE),

```

```

distribution.model = "sstd") # 偏t分布
ghyp_spec.mod= ugarchspec(
  variance.model = list(model = "realGARCH", garchOrder = c(1,1))
,
  mean.model = list(armaOrder = c(0,0), include.mean = FALSE),
  distribution.model = "ghyp") # 广义双曲线分布

Lst5=list(norm_spec.mod,std_spec.mod,sstd_spec.mod,ghyp_spec.mod)
#存放模型分布
AuLst=list() #存放 Au 拟合模型回归结果 1-4 是 RV

###Au 拟合模型参数估计
for (i in 1:4) {
  AuLst[[i]]=ugarchfit(data=Real_Au[, "rt"], spec = Lst5[[i]],
                        solver = "hybrid", realizedVol = Real_Au[
, "RV"])
}
names(AuLst)=c("AuRV_norm_sgarch.mod", "AuRV_std_sgarch.mod", "AuRV
_sstd_sgarch.mod", "AuRV_ghyp_sgarch.mod")

AuRVfit1=AuLst[["AuRV_norm_sgarch.mod"]]
AuRVfit1_par=AuRVfit1@fit$matcoef[, 1] #系数
AuRVfit1_p=AuRVfit1@fit$matcoef[, 4] #p 值
AuRVfit1_log=sum(-AuRVfit1@fit$log.likelihoods) #极大似然值
AuRVfit1=cbind(AuRVfit1_par,AuRVfit1_p) #合并

AuRVfit2=AuLst[["AuRV_std_sgarch.mod"]]
AuRVfit2_par=AuRVfit2@fit$matcoef[, 1] #系数
AuRVfit2_p=AuRVfit2@fit$matcoef[, 4] #p 值
AuRVfit2_log=sum(-AuRVfit2@fit$log.likelihoods) #极大似然值
AuRVfit2=cbind(AuRVfit2_par,AuRVfit2_p) #合并

AuRVfit3=AuLst[["AuRV_sstd_sgarch.mod"]]
AuRVfit3_par=AuRVfit3@fit$matcoef[, 1] #系数
AuRVfit3_p=AuRVfit3@fit$matcoef[, 4] #p 值
AuRVfit3_log=sum(-AuRVfit3@fit$log.likelihoods) #极大似然值
AuRVfit3=cbind(AuRVfit3_par,AuRVfit3_p) #合并

AuRVfit4=AuLst[["AuRV_ghyp_sgarch.mod"]]
AuRVfit4_par=AuRVfit4@fit$matcoef[, 1] #系数
AuRVfit4_p=AuRVfit4@fit$matcoef[, 4] #p 值
AuRVfit4_log=sum(-AuRVfit4@fit$log.likelihoods) #极大似然值
AuRVfit4=cbind(AuRVfit4_par,AuRVfit4_p) #合并

```

#似然函数值

```
AuRV_log=c(AuRVfit1_log,AuRVfit2_log,AuRVfit3_log,AuRVfit4_log)
```

```
AuRV_log #AuRV_ghyp 值最大
```

```
options(digits=3)
```

```
Au_estimate <- read.csv('./results/Auestimation.csv',header=T,sep  
=',',encoding = "ANSI")
```

```
colnames(Au_estimate) <- c("模型形式
```

```
", '$\\omega$', "$\\alpha$", "$\\beta$", '$\\tau_1$', '$\\tau_2$', '$\\delta$', '$\\lambda$', '$\\xi$', '$\\pi$')
```

```
AULOG=as.data.frame(t(AuRV_log))
```

```
colnames(AULOG)=c('正态分布','t 分布','偏 t 分布','广义双曲线分布')
```

```
rownames(AULOG)=c('对数似然函数值')
```

```
Auvar=read.csv("./results/AuVaR.csv",header=T,sep=',',encoding = "ANSI")
```

```
colnames(Auvar)=c("VaR 显著性水平","模型形式","LR 统计量","失败率","失败  
天数","P 值")
```

```
Auvar[2,1]="$\\alpha=0.01"
```

```
Auvar[-2,1]=""
```

```
total_Au = xts(read.zoo("./data/Real_Au.csv",header=TRUE,sep=',')  
)
```

```
best_model = ugarchfit(data=total_Au[, "rt"], spec = Lst5[[4]],  
solver = "hybrid", realizedVol = total_Au[,  
"RV"])
```

```
Auvar = best_model@fit[["sigma"]]^2
```

```
fit1=stats::arima(total_Au$rt,order = c(0,0,0))
```

```
Au_r=stats::residuals(fit1)
```

```
x=c("201908","10","12","202002","4","6","8","10","12")
```

```
xat=c(23,64,103,141,183,222,265,308,351)
```

```
plot(Auvar,type='l',xlab='日期',ylab='波动率',xaxt = "n")
```

```
lines(Au_r^2,type="l",xlab="日期",ylab="波动率  
",xaxt="n",col='grey')
```

```
axis(1,labels=x,at=xat)
```

```
legend("topleft",c("拟合波动率","实际波动率
```

```
"),lty=c(1,1),col=c("black","grey"),cex=1)
```