



基于希尔伯特-黄变换的 股指跳跃行为周期性特征研究

The Research of the Periodic Features of Stock Index Jumps Behavior Based on Hilbert-Huang Transformation

殷炼乾 李子祎

摘 要：股指波动率在金融时序的研究中具有核心作用，而其本质也是一种非线性、非平稳的信号。近年来提出的Hilbert-Huang Transform算法彻底摆脱了线性和平稳性的束缚，并有着清晰的物理含义。本文利用2012—2020年上证综指日波动率，以信号分解的角度对数据进行Hilbert-Huang变换。

关键词：希尔伯特-黄变换算法 经验模态分解 本质模态函数 平均震荡周期 波动率

Abstract: Stock index volatility has a key role in the research of the financial series, and its essence is a non-linear, non-stationary signals. The Hilbert-Huang Transform algorithm which proposed in recent years escape itself from the requirement of linear and smooth, and it has a clear physical meaning. The data of this article comes from the Shanghai Composite stock index diary volatility from 2012 to 2020, which is decomposed by Hilbert-Huang transform from the signal angle.

Keywords: Hilbert-Huang Algorithm, EMD, IMF, Average Oscillation Cycle, Volatility

作者单位：暨南大学国际商学院。本文为国家自然科学基金青年项目“任意模糊的复杂数据重建与跳跃检测方法研究”（11401094）和国家社科基金课题“数字货币的跨境反洗钱监管研究”（21BGL264）资助研究成果。



引言

作为金融市场研究的核心与关键，“波动率”是金融资产价格的变化程度，反映了金融资产的风险水平；其也是对资产收益率不确定性的测度，与投资收益紧密相连。以股票指数为例，“跳跃行为”是指股指收益率在日内近似连续时间内出现大幅波动的现象，其对于股指波动率的估计和预测有着十分重要的作用。因此，运用新型数据分析方法研究股指跳跃行为更深层次的特性，从而对股指波动率进行更精准的估计和预测，并进一步探寻提高证券市场收益的投资策略与更高效的风险管理方法，具有重要的现实意义。

目前对股指波动率的定量研究，国内外主要的方法均建立在一系列假设条件之上，偏向于运用传统时频分析方法研究线性、平稳、正态分布的金融时间序列。然而，真实的金融时间序列却更多体现了非平稳、非线性、波动率大的特性；此外，传统的时频分析技术，包括傅立叶变换和小波变换等，其本质均以傅立叶变换为理论依据，故在分析非平稳信号时，会出现假频等现象；小波变换还存在小波基的选择问题。因此，基于传统方法和技术的建模难以适用于具有非平稳、非线性特征的复杂系统（如证券市场价格等）的研究。鉴于当下研究困境，本文引入处理非线性、非平稳数据的新方法——希尔伯特-黄变换（Hilbert-Huang Transform, HHT），

对股指波动率进行研究。

HHT是通过基于瞬时频率的经验模态分解（Empirical Mode Decomposition，简称为EMD），先将信号分解为不同尺度的本质模态函数（Intrinsic Mode Function，简称为IMF），然后经过Hilbert谱分析将IMF转换为具有不同能量的边际频谱。与传统的信号处理方法相比，HHT彻底摆脱了线性和平稳性的束缚，并有着清晰的物理含义，能够得到信号的时间、频率、能量分布特征，而且是一种自适应性、适合突变信号的信号处理方法。时频分析的最终目的是要建立一种分布，以便能在时间和频率上同时表示信号的能量或者强度，使在时间域内难以观察到的信号特征在频率域内能十分清楚地显示出来，当下传统算法均存在一定缺陷，而HHT算法可同时在时域和频域两方面达到高精度，是实现该目的较好方法。因此，如果能通过对上证综指波动率中的已实现波动率、跳跃波动率、双幂次变差等信号进行Hilbert-Huang变换，得到时频分布后就可以对各种信号进行综合、分析、对比、处理，并提取信号中所包含的特征信息。

本文主要研究的信号序列是以“已实现波动率”理论为基础，运用列维（Levy）分离定理，得到跳跃行为的非参数估计量。在跳跃行为研究方面，Andersen & Bollerslev（2003）以等间隔时间序列为基础，证明了Realized Volatility



是积分波动的无偏估计,提出了全新的波动率估计方法。Barndorff-Nielsen & Shephard (2006)提出了“已实现”二次幂变差的方法,可用于检验跳跃是否存在,首次实现了对跳跃行为的直接测度。随着高频金融数据的广泛应用,Andersen et al. (2006)基于Barndorff-Nielsen & Shephard (2006)二次幂变差(Bi-Power Variation)理论,使用非参数方法将“已实现”波动率分解为连续样本路径方差(Continuous Sample Path Variation)和离散跳跃方差(Discontinuous Jump Variation),实现了高频数据中跳跃行为的剥离,并对资产的日内跳跃行为进行了实时检验。胡志军(2013)改进了Andersen et al. (2010)的资产价格日内序列跳跃检验方法,对中国股市资产价格跳跃行为进行了细致地刻画。殷炼乾等(2015)用基于高频数据集的分析方法对中国资产价格跳跃行为进行了详细研究。非参数方法属于直接研究范畴,其研究思想是根据资产价格行为所具备的直观特征来构造统计量,将跳跃行为引起资产价格波动部分剥离出来,直接测度跳跃行为。

在信号处理研究方面,丁志宏和谢国权(2009)针对小波变换的不足之处,运用EMD对沪深300指数日收益率时序进行多时间尺度分解,发现其波动存在准2日、准4—5日、准15日、准28日、准70日、准140—190日及准240日等波动周

期,并分析了各分量的趋势变化,对金融时间序列多分辨率进行了实证研究。王磊(2008)利用Hilbert-Huang变换能在时频两域都达到高精度的优势,通过Hilbert-Huang变换和积分获得边际谱分布,并对各个不同频率的能量分布进行尖峰分析,总结出2005年10月11日至2007年10月5日这500个交易日内,每日收盘价的隐藏波动周期约为164日。滕飞和董小刚(2008)提出了一种基于Hilbert-Huang变换的信号周期性分析方法,通过分析信号的非线性对主频分布的影响,找到了具有丰富高频含量的近似周期信号的主频与其周期性的近似对应关系。于文婷(2014)对沪深300股指期货2010年4月至2013年12月数据序列进行EMD分解以及Hilbert谱、边际谱的计算,以此观察分析沪深300股指期货的周期性特征,并提出基于Hilbert-Huang变换的布林通道交易策略。以EMD为基础的Hilbert-Huang变换不可否认地丰富了信号分解在股指波动率的应用。

本研究也是从信号分解的角度,选取比较典型的已实现波动率(RV)、跳跃波动率(JV)、二次幂变差(BV),对几种波动率的波动性和周期性进行探究,并对瞬时振幅、瞬时频率等进行刻画和分析。结果表明:股指波动率经过Hilbert-Huang变换处理和对比下,周期性和波动性在不同的尺度上均存在一定的关系,各种Hilbert谱在频域上的物理意义也得到了



解释，为Hilbert-Huang变换在股指波动率方面的实证研究提供了参考材料。

跳跃信号与Hilbert-Huang变换

波动率

Merton (1976) 指出当第 t 日的交易时间划分的区间数 m 趋于无穷大时，资产对数价格日内收益率的平方和的已实现波动率 (Realized Volatility) RV会概率一致均匀的收敛于二次方差QV。基于Merton的研究思想，Alizadeh et al. (2002) 等人进一步发展得出了已实现波动率RV的渐进理论，即：

$$RV_t = \sum_{i=1}^m r_{t,i}^2 \xrightarrow{\Delta \rightarrow 0, p} \int_{t-1}^t \sigma_s^2 dW_s + \sum_{t-1}^t \kappa_s^2 \quad (1)$$

其中 $r_{t,i}^2$ 为日内各区间收益率 $r_{t,i}$ ($i=1,2,\dots,m$)的平方序列。当 m 趋于无穷大时，即每个区间的时间间隔 Δ 趋于零时，已实现波动率RV会依概率一致均匀的收敛于第 t 日二次方差QV。已实现双幂次变差实际上是资产在第 t 日内前后相邻两收益率的绝对值之积，用公式表示为：

$$BV_t = \frac{\pi}{2} \frac{m}{m-1} \sum_{i=2}^m |r_{t,i}| |r_{t,i-1}| \xrightarrow{\Delta \rightarrow 0, p} \int_{t-1}^t \sigma_s^2 dW_s \quad (2)$$

其中 $r_{t,i}^2$ 为第 t 天的日内各区间收益率 $r_{t,i}$ ($i=1,2,\dots,m$)的平方序列， $r_{t,i-1}$ 为第 $t-1$ 天的日内收益率。当 m 趋于无穷大时，即每个区间的时间间隔 Δ 趋于零时，已实现双幂次变差BV会依概率一致均匀的收敛于第 t 天的积分方差IV。

综合式 (1) 和 (2)，可以得到已实现波动率RV与已实现双幂次变差BV之差的依概率一致均匀收敛于第 t 天的跳跃方差JV的估计量，即

$$RV_t - BV_t \xrightarrow{\Delta \rightarrow 0, p} JV_t = \sum_{t-1}^t \kappa_s^2 \quad (3)$$

至此，我们实现了对跳跃行为的直接测度，跳跃方差就是资产价格跳跃行为导致资产波动的部分。因此跳跃成分所引起的波动性部分可以定义为实现跳跃变差：

$$JV_t = I_{\{ZJ_t > \Phi_{1-\alpha}\}} (RV_t - BV_t) \quad (4)$$

其中， $I_{\{\cdot\}}$ 表示指示函数，若 $ZJ_t > \Phi_{1-\alpha}$ 则 $I_{\{\cdot\}} = 1$ ，否则为0； $\Phi_{1-\alpha}$ 表示标准正态分布的 $1-\alpha$ 分位数， α 是估计 JV_t 时所选择的置信水平。

经验模态分解EMD

经验模态分解的本质是一个筛选过程，其利用拟合得到的上下包络线的平均值，确定出“瞬时平衡位置”，进而提取固有模态函数IMF。为了能确定IMF，其必须满足以下两个条件：第一，信号的极值点（极大值或极小值）数目和过零点数目相等或最多相差一个；第二，由局部极大值构成的上包络线和由局部极小值构成的下包络线的平均值为零。而经验模态分解的基本过程可概括如下：

寻找原信号JV (t) 所有的极大值点，并用三次样条插值函数拟合形成原数据的上包络线。同理，找出所有的极小值



点，并将所有的极小值点通过三次样条插值函数拟合形成数据的下包络线；

计算上、下包络线的平均值，记作 $m1$ ，将原数据序列 $JV(t)$ 减去该平均包络 $m1$ ，得到一个新的数据序列 $h1$ ，即 $X(t) - m1 = h1$ ；

如果 $h1$ 还存在负的局部极大值和正的局部极小值，说明这还不是一个本质模态函数，需要进行“筛选”，重复上述步骤，继续得到 $h2, h3 \dots$ 如果有 $h(t)$ 满足固有模态函数的两个条件，此时就将得到的 $h(t)$ 作为一个 IMF1，进行下一步；

将原始信号 $JV(t)$ 减去 IMF1，剩余的信号作为原始信号继续从第1步开始，以此求取剩余的 IMF (t) ；

最后得到不可分解的信号，如单调序列或常值序列条件时，终止筛选过程。剩余的信号 $e(t)$ 代表了信号的均值、趋势项 (RES)。由此可知原始序列是由多个 IMF 与一个 RES 组成。如上所述，整个过程就像筛选过程，根据时间特性把本质模态函数从信号中提取出来。需要注意的是，根据上文，提取的 IMF 必须满足两个条件，但是在实际筛选过程中，严格满足这两个条件的信号几乎不存在。因此如果以这两个条件为依据判定 IMF，可能得不到结果或要以冗长的程序执行时间为代价。为了保证 IMF 的调幅和调频都具有物理意义，同时考虑到程序可行性，必须制定一个结束筛选的准则。习惯上，可以通

过标准偏差 SD 来完成，通过实际情况需要的精确度，来控制筛选过程的标准偏差 SD 。

Hilbert谱分析

时域分析主要关心信号的频谱随时间变化的情况。瞬时频率表征了信号在局部时间点上的瞬态频率特性，整个持续期上的瞬时频率反映了信号频率的时变规律。对于跳跃时间序列 $JV(t)$ ，对其进行 Hilbert 变换，可以得到 $Y(t)$ 如下：

$$Y(t) = \frac{1}{\pi} PV \left(\int_{-\infty}^{\infty} \frac{X(\tau)}{t - \tau} d\tau \right) \quad (5)$$

其中 PV 为柯西主值 (Cauchy principal value)。该式表示 $Y(t)$ 是 $X(t)$ 与 $1/\pi t$ 的卷积。通过这个定义， $X(t)$ 和 $Y(t)$ 组成了一个共轭复数对，于是可以得到一个解析信号 $Z(t)$ 如下：

$$Z(t) = X(t) + iY(t) = a(t)e^{i\theta(t)} \quad (6)$$

$$\text{其中} \begin{cases} a(t) = [X^2(t) + Y^2(t)]^{1/2} \\ \theta(t) = \arctan \left(\frac{Y(t)}{X(t)} \right) \end{cases} \quad (7)$$

从理论上讲，虚部的定义方法有很多种。但是 Hilbert 变换为其提供了一个唯一的虚部值，这就使得其结果成为一个解析函数。瞬时频率是相位导数，因此得到了相位，就可得到瞬时频率：

$$\omega = \frac{d\theta(t)}{dt} \quad (8)$$

由希尔伯特变换得出的振幅和频率



都是时间的函数，如果用三维图形表达幅值、频率和时间之间的关系，或者把振幅用灰度的形式显示在“频率-时间”平面上，就可得到Hilbert谱 $H(w, t)$ 。如果把 $H(w, t)$ 对时间积分，就得到希尔伯特边际谱 $h(w)$ ：

$$h(w) = \int_0^T H(w, t) dt \quad (9)$$

边际谱提供了对每个频率的总振幅的量测，表达了整个时间长度内累积的振幅。如果振幅的平方对时间积分，可以得到希尔伯特能量谱：

$$ES(w) = \int_0^T H^2(w, t) dt \quad (10)$$

希尔伯特能量谱提供了对于每个频率的能量的量测，表达了每个频率在整个时间长度内所累积的能量。

实证分析

EMD信号分解过程

以2020年上证综指的已实现波动率、双幂次变差、跳跃波动率的数据为例，经过EMD分解为各个尺度的IMF分量，图1各图从左到右、从上到下依次为原信号、IMF1—IMF6及残差项RES。

从波动的细节上看，IMF1分量相比其他IMF分量，波动最剧烈，表现出大幅度的高频震荡，其波动细节与原信号最接近，所以其代表的是最高频信号，很好地保留了原信号的细节部分；IMF2—IMF3

分量仍处于波荡之中，其相对于IMF1频率稍低，也表现了高频特性；IMF4—IMF6，波动信息明显开始减少，只有在原信号表现为突变的情况，才保留了波动的细节，更多表现的是一种低频的趋势；残差项RES的趋势吻合原信号，解释了波动率总体下降和上升的趋势，并且RES中最低谷的位置，对应的就是原信号波动率最低的时候，说明了RES保留了较多的能量，足以影响原信号。在数量关系上，已实现波动率等于双幂次变差与跳跃波动率之和。对比三种波动率的各个尺度的IMF可以发现，双幂次变差和已实现波动率除了一些细节，信号基本相同。在IMF6和RES中可以看出，跳跃波动率和双幂次变差、已实现波动率的趋势基本相反。在其他尺度上，跳跃波动率大的时候，双幂次变差、已实现波动率却比较小；跳跃波动率小的时候，双幂次变差、已实现波动率却比较大。

探索各个IMF的平均震荡周期

为了有足够的数量，以2012—2019年共8年的波动率日数据为例，计算各个IMF的平均震荡周期，此时定义平均震荡周期=总天数/（极大值天数+极小值天数）/2。结果见表1，可得到以下结论。

从IMF1~IMF10，分解尺度越大，极值点越少，平均震荡周期越长。分解尺度越大的分量，代表的越是低频的变动，所以也就是一个长期性的变动；

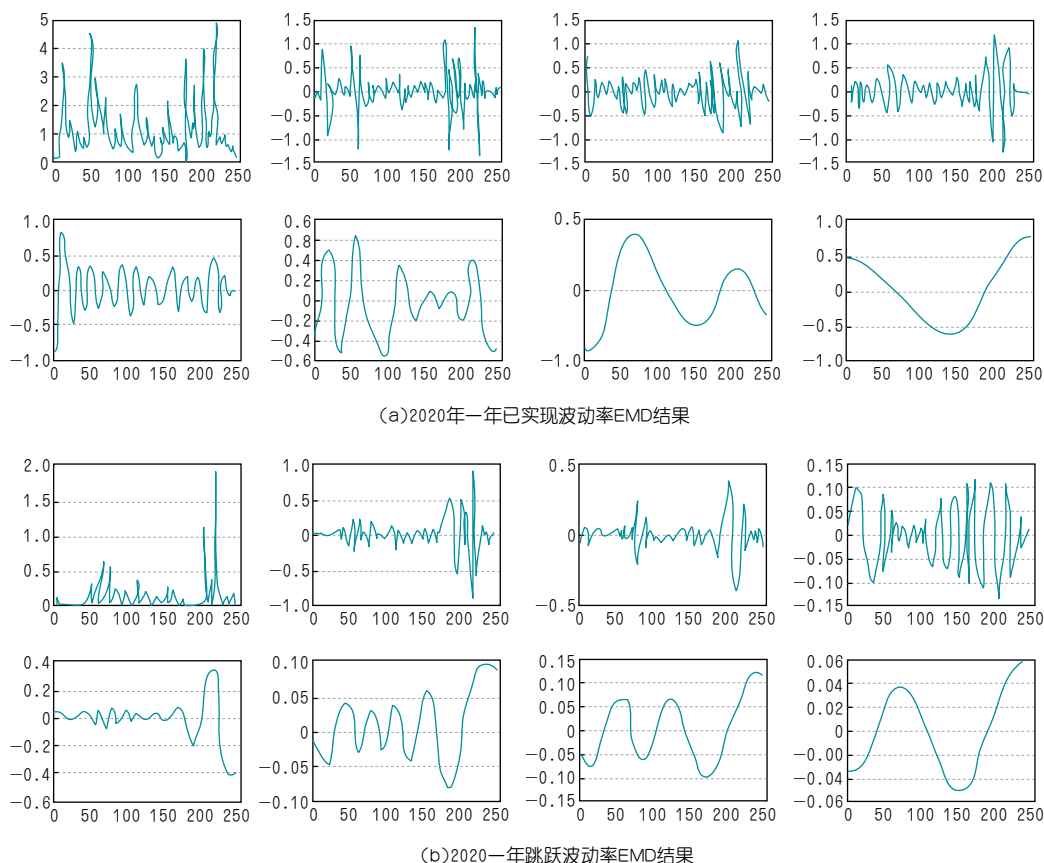


图 1 波动率 EMD 结果

从表上的天数大概可以看出，对于已实现波动率来说，IMF1分量信号大致代表的是3天的频率变动，IMF2代表的是一个星期的平均变动趋势。而到了IMF10，因为数据量不够多，已经没有实际意义了。其具体代表了天数还会因为EMD本身求极值点时固有的一些边界问题等以及采用的数据量不够等原因而出现误差，因此仅供

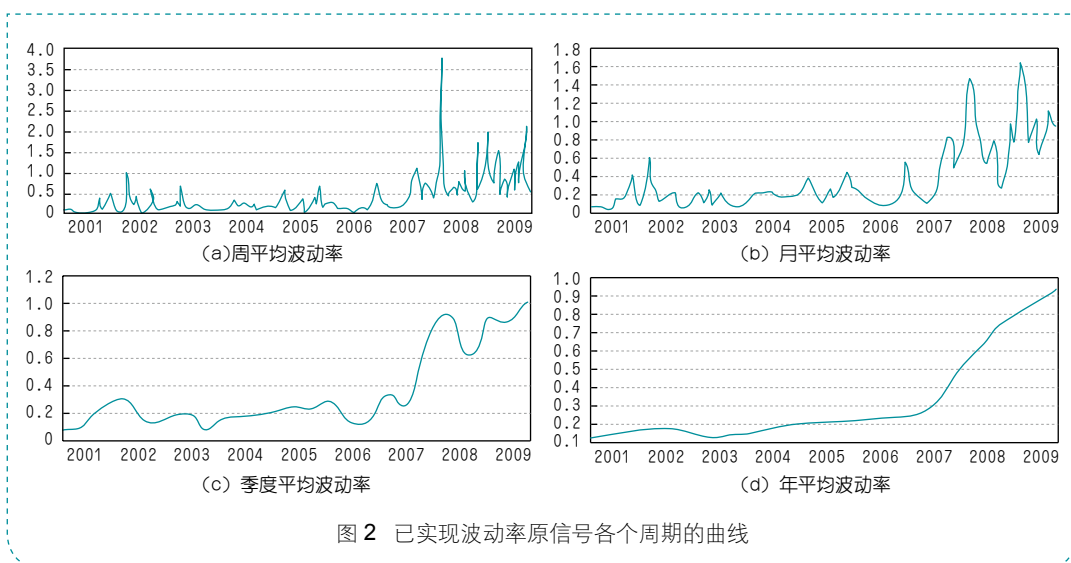
参考；

虽然总天数一样，但是对比三种波动率的分解结果，跳跃波动率多一个尺度。对比极值的天数和平均震荡周期可以发现：在高频阶段，由于跳跃波动率的数据密度比较小，大部分是0。因此，其极值点比较少。所以，已实现波动率（RV）的极值天数>双幂次变差（BV）的极值天数



表 1 IMF 的平均震荡周期

总天数1916天	极大值的天数			极小值的天数			平均震荡周期		
信号	RV	BV	JV	RV	BV	JV	RV	BV	JV
原始信号	584	701	161	584	701	22	3.280	2.733	6.818
IMF1	635	626	281	635	625	281	3.017	3.063	10.164
IMF2	359	340	188	360	341	189	5.330	5.627	15.027
IMF3	209	183	128	209	183	127	9.167	10.469	23.801
IMF4	129	102	81	129	102	80	14.853	18.784	45.082
IMF5	65	52	43	64	52	42	29.705	36.846	66.069
IMF6	36	32	29	36	32	29	53.222	59.875	132.138
IMF7	19	14	14	18	14	15	103.568	136.857	147.385
IMF8	8	7	13	9	8	13	225.411	255	255.467
IMF9	3	2	8	4	3	7	547.429	766.4	319.333
IMF10	2	1	6	1	1	6	1277.33	1916	766.4



>跳跃波动率（JV）的极值天数，所以跳跃波动率的平均震荡周期最大。而在IMF7之后，由于其总共的分解尺度比其它波动率多一次，因此其低频的缓和程度又不如双幂次变差，因此其波动周期又比较短，

跳跃波动率（JV）的极值天数逐渐增加。

探索不同波动周期的平均波动率与重构后的IMF的关系

不同波动周期的平均波动率

上文探索了不同IMF的震荡周期，而



对于股指波动率来说，也有其自身的周期，如以周为周期的波动、以月为周期的波动，周期越长，波动就越缓和，越接近低频的特性。那么股指本身的波动周期是否会与IMF有关呢？首先求各个周期的平均指数，得到直观的波动率图如图2（限于篇幅，仅列示已实现波动率）。

不同尺度的IMF进行重构

每个IMF仅代表了那一个分量的波动，而股指代表了原信号。因此，需要通过将IMF进行从残差、低频、到高频逐步相加进行重构，然后再将重构之后的信号的周期与股指的周期进行比较。图3从左到右、从上到下分别是已实现波动率的不同尺度IMF重构信号。

探究每个重构IMF代表的波动周期并得出数量关系

同样，对跳跃波动率也进行如上探索（限于篇幅，图未列示），对比两种不同的波动率发现，对于探究每个重构的IMF的最佳解释周期，已实现波动率的意义比跳跃波动率的更大。这是因为跳跃波动率重构之后的IMF的负值太大，而原信号全部都是正值，还有由于在前期跳跃波动率过于低频，基本不可能吻合。而已实现波动率除了一些特别大的负值，基本都吻合。通过上述同样的方法，不断缩小日期范围，确定每个重构的IMF与多长时间作为周期的波动率最吻合，得到以下结论（限于篇幅，表2仅列示跳跃波动率相关结果）。

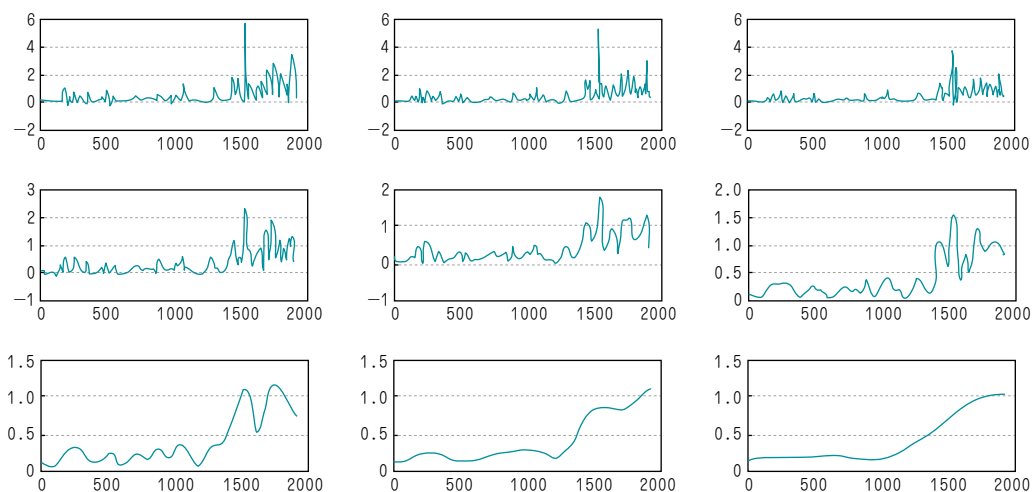


图3 已实现波动率的IMF2_c、...、IMF10_c



(1) 已知分解尺度越大的IMF，曲线越平缓，代表的波动周期就越长，因此重构之后所代表的周期就越来越长；

(2) 无论是已实现波动率还是跳跃波动率，各个尺度的IMF与重构后的IMF的平均震荡周期都基本相等（当分解尺度小于IMF6的时候），或重构后的略大于原IMF。这说明重构后，极值点略有所减少，图形变得平和一点，比较少负值，并且更接近于原信号，所以周期略长；

(3) 由于各个IMF代表的是不同频段的波动周期，用同样的方法，并不能推测出每个IMF所代表的原信号的波动周期。而重构后的IMF经过缓和处理，更接近于原信号，并且能推测出各个重构后的IMF所代表的原信号的波动周期；

(4) 重构之后的IMF会更高频，因为其吸收了好几个IMF的细节。对于已实现波动率，重构之后的IMF所代表的天数会

比单独的IMF短，大致为2~3倍的关系，而跳跃波动率则是4~5倍的关系；

(5) 跳跃波动率重构后的IMF所能代表的原信号的波动周期的精确度下降了不少，因为前期基本不吻合，只有当后期高频的时候才会和趋势基本一致。

探索对不同尺度的IMF与原始信号的相关性，方差贡献率

由表3，从各个IMF与原始信号的相关性可以看出：对于已实现波动率，RES与原始信号的相关性最高，说明RES代表了已实现波动率总体的趋势。而对于跳跃波动率，与已实现波动率不同，IMF1与原始信号的相关性最高。这是因为跳跃波动率是一个跳变性的波动率，当IMF1频率变化非常大，细节占的比重非常大时，其与原信号的相关性、方差贡献率都是非常大的。而双幂次变差作为一种普通的信号，RES代表了已实现波动率总体的趋势，这

表 2 跳跃波动率各个重构 IMF 代表的波动周期及其各个 IMF 的平均震荡周期

重构后的IMF	代表波动周期	平均震荡周期	各个尺度的IMF	平均震荡周期
IMF2_c	2	10.413	IMF2	10.164
IMF3_c	3	15.514	IMF3	15.027
IMF4_c	4	24.408	IMF4	23.801
IMF5_c	6	43.545	IMF5	45.082
IMF6_c	18	67.228	IMF6	66.069
IMF7_c	30	123.613	IMF7	132.138
IMF8_c	40	166.609	IMF8	147.385
IMF9_c	60	273.714	IMF9	255.467
IMF10_c	240	425.778	IMF10	319.333
IMF11_c		958	IMF11	766.4

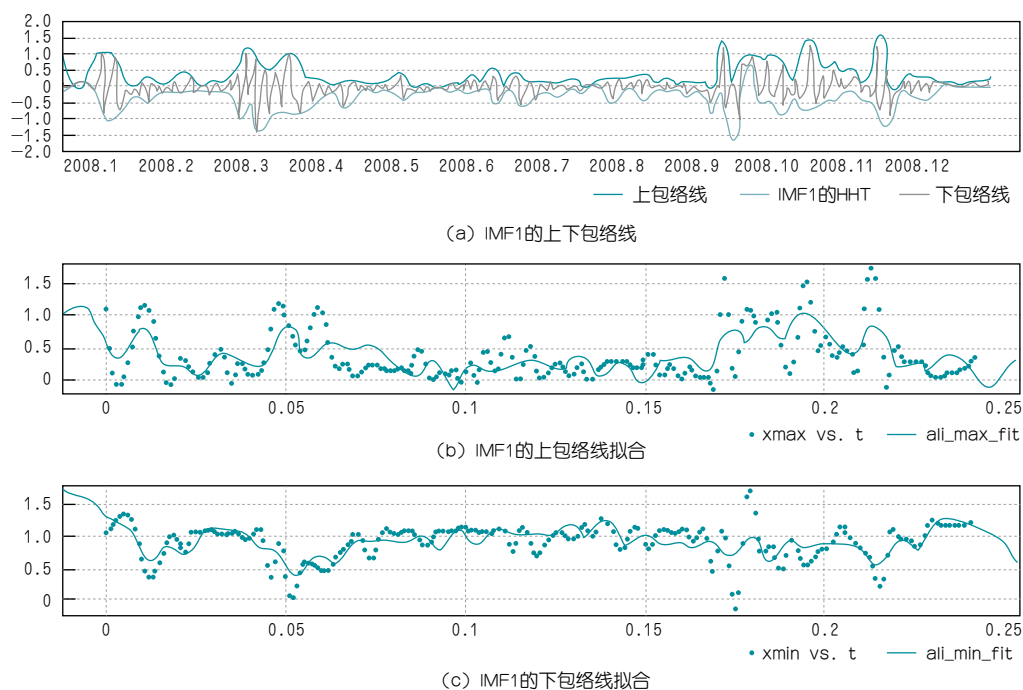


图 4 Hilbert 变换后 IMF1 包络线的拟合情况

表 3 各个 IMF 与原始信号的相关性，方差与方差贡献率

统计量	各IMF分量及最后剩余分量 RES与原始信号的相关性		各IMF分量的方差：为平方的 均值减去均值的平方		各IMF分量的方差百分比，也 就是方差贡献率	
波动率类型	已实现波动率	跳跃波动率	已实现波动率	跳跃波动率	已实现波动率	跳跃波动率
IMF1	0.309	0.3696	0.050	0.0064	14.044	44.1415
IMF2	0.212	0.2513	0.026	0.0042	7.335	29.1022
IMF3	0.180	0.3203	0.037	0.0010	10.329	6.8653
IMF4	0.319	0.2431	0.058	0.0014	16.340	9.5596
IMF5	0.265	0.1189	0.027	0.0004	7.462	2.5049
IMF6	0.147	0.0676	0.012	0.0002	3.274	1.2566
IMF7	0.262	0.1166	0.020	0.0003	5.592	2.0169
IMF8	0.193	0.1330	0.015	0.0002	4.158	1.2267
IMF9	-0.003	0.0660	0.006	0.0001	1.689	0.5801
IMF10	0.224	0.0274	0.026	0.0000	7.317	0.3023
RES/IMF11	0.474	0.1408	0.080	0.0001	22.458	0.3902
RES		0.2026		0.0003		2.0536



是一个正常现象。从方差与方差贡献率来看：对于已实现波动率，IMF1的波动性最大，代表了最高频的分量，所以其方差也是最大，方差贡献率也是最大。而跳跃波动率，其方差都降了两个数量级，应该是信号在各个IMF的分配没有已实现波动率那么平均。

基于Hilbert变换的波动率信号处理 对Hilbert变换后的包络线进行拟合

采用2020年一年的数据，在MATLAB软件中，通过寻找极值点对各个IMF画出上包络线和下包络线，可以发现，由于边界上的点不可能同时是极大值和极小值，因此会出现边界效应。在MATLAB中通过其自带的carve fitting tool进行拟合，下图为双幂次变差的IMF1经过Hilbert变换后得到的上包络线和下包络线的拟合情况：

通过拟合发现，无论是已实现波动率还是跳跃波动率，每一个IMF的最佳拟合分布都是sum of sine，8阶的时候拟合效果最佳，因此可以说，用sum of sine 分布来拟合IMF是较好的选择，具体拟合效果如表4所示。通过观察实际的拟合线（限于篇幅未列示）发现：低频的时候，其包络情况本身就因为算法固有的弊端，如端点效应等而不可信；而高频的时候下包络线效果不好是因为一些数据中有一些非常小的极值点而影响了下包络线的拟合效果。所以，按照数据有效性来说，总体情况都是下包络线的效果会好于上包络线。因此在实际情况中，中频下包络线拟合的参考性会更好，并且包络线最佳的拟合分布的sum of sine（8），在IMF5的时候拟合效果好，在IMF6的时候又开始反复。

表 4 波动率各个 IMF 的上下包络线的拟合效果

Sum of sine(8)	已实现波动率				跳跃波动率			
	SSE:	R方	调整R方	RMSE:	SSE:	R方	调整R方	RMSE:
a1i_max_fit	17.68	0.4753	0.4202	0.2842	2.511	0.8082	0.7882	0.1068
a1i_min_fit	14.11	0.5087	0.4571	0.2538	3.568	0.6689	0.6343	0.1273
a2i_max_fit	0.6093	0.9725	0.9696	0.05275	0.04748	0.9839	0.9822	0.01469
a2i_min_fit	0.7285	0.9046	0.9562	0.05768	0.6019	0.8352	0.818	0.5023
a3i_max_fit	0.1064	0.9955	0.9951	0.02204	0.004631	0.9888	0.9877	0.004588
a3i_min_fit	0.09523	0.9966	0.9963	0.02085	0.000881	0.9974	0.9971	0.002001
a4i_max_fit	0.3274	0.9797	0.9776	0.03867	0.00834	0.9989	0.9988	0.006157
a4i_min_fit	0.05091	0.9953	0.9948	0.01525	1.902	0.968	0.9646	0.09298
a5i_max_fit	0.005239	0.9996	0.9995	0.00489	0.000480	0.9999	0.9999	0.001478
a5i_min_fit	0.000687	0.9999	0.9999	0.00177	0.001518	0.9959	0.9955	0.002627
a6i_max_fit	0.04702	0.9964	0.996	0.01465	0.005954	0.9968	0.9965	0.005202
a6i_min_fit	0.426	0.9825	0.9807	0.0441	0.01018	0.9882	0.9869	0.006802

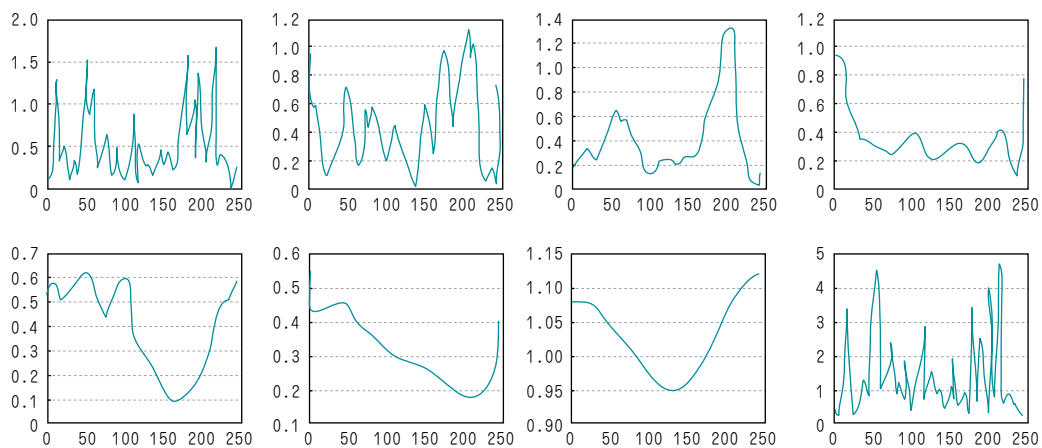


图 5 已实现波动率各个 IMF 经过 Hilbert 变换得到的时间 - 振幅图

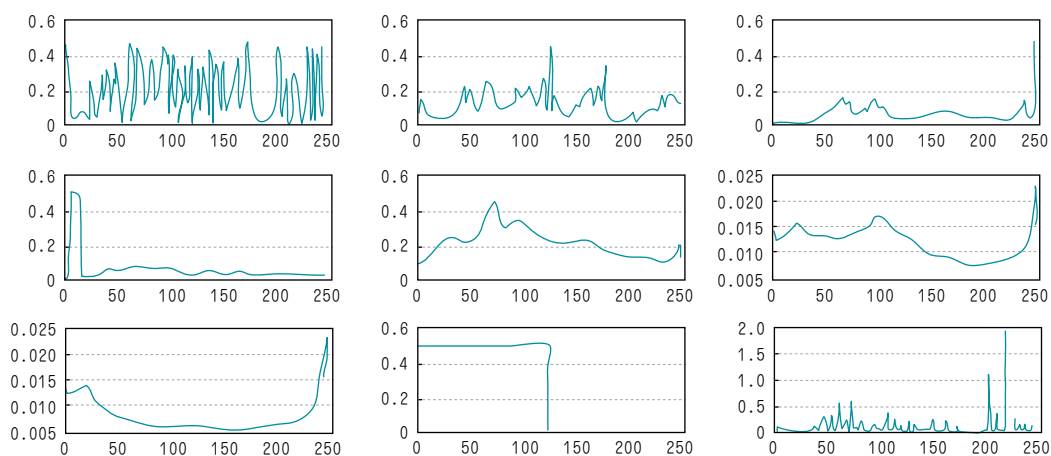


图 6 跳跃波动率经过 EMD 分解后再通过 Hilbert 变换得到瞬时频率

Hilbert变换得到“时间-振幅”曲线

Hilbert变换得到“时间-振幅”曲线如图5所示（仅列示已实现波动率）。振幅曲线与原IMF信号类似，分解尺度小的，其代表的即是高频的分量，因此其振幅变

化就快。如IMF1—IMF3代表了原始信号振幅变化快的细节部分的位置。由于各个IMF的极大值代表了在那一个尺度上的突变情况，而高频分量的突变更容易被保留下来，所以如果在最高频的前三个IMF找



那个都出现的极大点位置，就是突变点的位置（极大值并不一定是突变点，只有突然上升到极大值点才可以算作突变点）。从图中可以看到，在前面3个IMF的接近200的地方，都出现了极大值，这也就是原信号（最右下角）的一个突变点。

Hilbert变换得到“时间-频率”曲线

Hilbert变换得到的“时间-频率”曲线如图6所示（仅列示跳跃波动率）。频率的概念是从机械旋转运动来的，定义为角速度，对于周期运动，角速度即角频率。通常以反时针为正，因此转动的正频率是反时针旋转角速度，负频率就是顺时针旋转角速度。这就是它的物理意义，正、负号不影响它的物理意义。但是通常由于Hilbert变换只是单纯地求信号的瞬时振幅，频率和相位，有可能出现负频率，对于研究股指波动率来说，可能会没有意义。若在Hilbert变换前，先将信号进行EMD分解，得到各个不同尺度的分量，再对每一个分量进行Hilbert变换则可获得有实际意义的瞬时频率。

对于“时间-频率”图，竖轴展示了在不同的频率上，波动率变化的快慢。通过实际观察对比“时间-频率”图（限于篇幅未全部列示）发现：已实现波动率，从IMF1—IMF6，在高频的IMF1分量中因为细节不断变化，即高频上变化快，所以瞬时频率都是最大的，到了IMF5分量，此时在低频上呈现下降趋势，无太大改变，频

率出现了非常低频的情况；跳跃波动率也是一样，在低频上变化比较慢，所以频率比较低。而对于两种波动率的RES，都出现了频率截断的情况，即从高频到基本0频率，这是因为RES代表的是趋势，频率变化快的部分则是高频，频率变化慢的部分则是低频。如跳跃波动率前期虽然数值不大，但变化快，所以高频；而后期变化慢，则基本0频率。已实现波动率也是，中间变化快，RES的频率就比较大。

结论

长期以来，证券市场波动性的研究始终是金融市场研究的核心问题，关于股指波动率的研究也涉及到多方面的理论、方法、研究和技术，其中波动率的性质相关研究更是数不胜数。然而，证券市场作为一个复杂的系统，金融资产价格并非一个简单的时间序列，而是具有非平稳、非线性特征的复杂信号，需要涉及更前沿的分析处理方法。

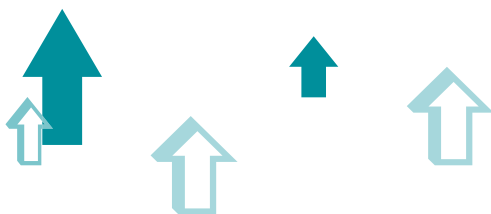
本文针对传统时间序列法在处理具有非平稳非线性特征的复杂系统时间序列时的缺陷，采用Hilbert-Huang Transform算法，从信号分解的角度处理股指波动率，选取比较典型的已实现波动率（RV）、跳跃波动率（JV）、双幂次变差（BV），以此算法的流程为框架，完成了对这几种波动率的分析 and 比较，也得出了一些在信号分解角度特有的性质，如通过计算信号



的极值点、形状和吻合度，来推测各种周期，并得到数量关系。然而，Hilbert-Huang Transform算法目前主要在医学（探测心率、血压及传染病扩散）、交通（探测公路桥梁安全）、地理（地震工程）、航天（卫星信号）等自然科学领域广泛应用，国内外在其理论和算法方面的研究相对较少，在金融领域的运用甚少。本文利用该方法对股指时间序列进行分析处理，还有一些问题需要解决，如数据有限，没有进行多次推算检验结论以及没有考虑

Hilbert-Huang Transform算法的端点效应等带来的影响。因此在实际应用中，如需参考本文的思路，或在此基础上进行趋势分析并提出具体对策，还需继续积累和完善。^[N]

学术编辑：曾一已



参考文献

- [1]丁志宏,谢国权.金融时间序列多分辨率实证研究的EMD方法[J].经济研究导刊,2009(06):61-63.
- [2]胡志军,沈根祥.中国股市资产的价格跳跃行为——基于上证综指、深证成指的高频数据研究[J].上海经济研究,2013,25(04):88-99.DOI:10.19626/j.cnki.cn31-1163/f.2013.04.008.
- [3]滕飞,董小刚.基于Hilbert-Huang变换的金融数据周期性分析及应用[J].长春工业大学学报(自然科学版),2008(03):288-291.
- [4]邵锡栋,殷炼乾.基于实现极差和实现波动率的中国金融市场风险测度研究[J].金融研究,2008(06):109-121.
- [5]殷炼乾,周雨田,吴华妹.中国资产价格跳跃行为研究--基于高频数据集的分析[J].统计与信息论坛,2015(5):57-62.
- [6]于文婷.基于Hilbert-Huang变换的布林通道交易策略[D].浙江大学,2014.
- [7]王磊.HHT与AC算法在金融数据分析中的应用研究[D].浙江大学理学院浙江大学,2008.
- [8]Alizadeh S, Brandt M W, Diebold F X. Range based estimation of stochastic volatility models[J]. The Journal of Finance, 2002, 57(3): 1047-1091.
- [9]Andersen T G, Bollerslev T, Diebold F X, et al. Modeling and forecasting realized volatility[J]. Econometrica, 2003, 71(2): 579-625.
- [10]Andersen T G, Bollerslev T, Frederiksen P H, et al. [Realized Variance and Market Microstructure Noise]: Comment[J]. Journal of Business & Economic Statistics, 2006, 24(2): 173-179.
- [11]Andersen T G, Bollerslev T, Frederiksen P, et al. Continuous time models, realized volatilities, and testable distributional implications for daily stock returns[J]. Journal of Applied Econometrics, 2010, 25(2): 233-261.
- [12]Barndorff-Nielsen O E, Shephard N. Impact of jumps on returns and realised variances: econometric analysis of time-deformed Lévy processes[J]. Journal of econometrics, 2006, 131(1-2): 217-252.
- [13]Merton R.C.,Option pricing when underlying stock returns are discontinuous[J], Journal Financial Economics,1976,3(1):125-144.