

## 中证指数研究报告

# 实际波动率指数



## 目 录

一、波动率的概述 .....	2
(一) 何为波动率 .....	2
(二) 波动率的类型 .....	2
(三) 波动率的度量 .....	2
二、实际波动率与 VIX 指数区别 .....	3
(一) VIX 指数 .....	3
(二) 实际波动率指数 .....	4
三、实际波动率指数计算 .....	4
(一) 波动率序列 .....	4
(二) 小波分解 .....	5
(三) 非线性动力系统 .....	6
(四) 最大 Lyapunov 指数计算 .....	8
(五) 实际波动率指数生成 .....	10
四、结论 .....	11

## 图表目录

图表 1 波动率度量方法 .....	3
图表 2 实际波动率指数计算流程 .....	4
图表 3 沪深 300 波动率分解 .....	6
图表 4 延迟时间确定 .....	7
图表 5 最佳嵌入维数确定 .....	8
图表 6 最大李雅普诺夫指数计算 .....	10
图表 7 实际波动率指数表现 .....	11

波动率是金融领域非常重要的变量，组合管理、资产定价、风险控制等都离不开对波动率的准确度量。而且，波动率本身也具有重要的投资价值，如 VIX 指数，不仅成为全球投资者评估美国股票市场风险的主要依据之一，而且以 VIX 指数为标的开发的期货期权产品也广受市场的青睐。

## 一、波动率的概述

### （一）何为波动率

一般来说，波动率是一个市场或一种证券表现在一段短时间内大幅上升或下跌的统计性指标。通常波动率有多种定义，如标准差 (Standard deviation)，价格平方差 (Difference in prices squared)，收益平方 (Return Squared)，价格偏差绝对值 (Absolute value of the difference in prices)，收益绝对值 (Absolute value of the returns)。

### （二）波动率的类型

通常，波动率可以分为以下四类：

#### 1、实际波动率

实际波动率是指对有效期内投资回报率波动程度的度量，由于投资回报率是一个随机过程，实际波动率永远是一个未知数。

#### 2、历史波动率

历史波动率是指投资回报率在过去一段时间内所表现出的波动率，它由标的资产市场价格过去一段时间的历史数据反映。

#### 3、预测波动率

预测波动率是指运用统计推断方法对实际波动率进行预测得到的结果，并将其用于期权定价模型，确定出期权的理论价值。因此，预测波动率是对期权进行理论定价时实际使用的波动率。预测波动率并不等于历史波动率，前者是对实际波动率的理解和认识，历史波动率往往是这种理论和认识的基础。

#### 4、隐含波动率

隐含波动率是期权市场投资者在进行期权交易时对实际波动率的认知。由于期权定价模型给出了期权价格与五个基本参数  $S_t, X, r, T-t, \sigma$  之间的定量关系，通过前 4 个基本参数和期权的市场价格可以解出隐含波动率  $\sigma$ 。

### （三）波动率的度量

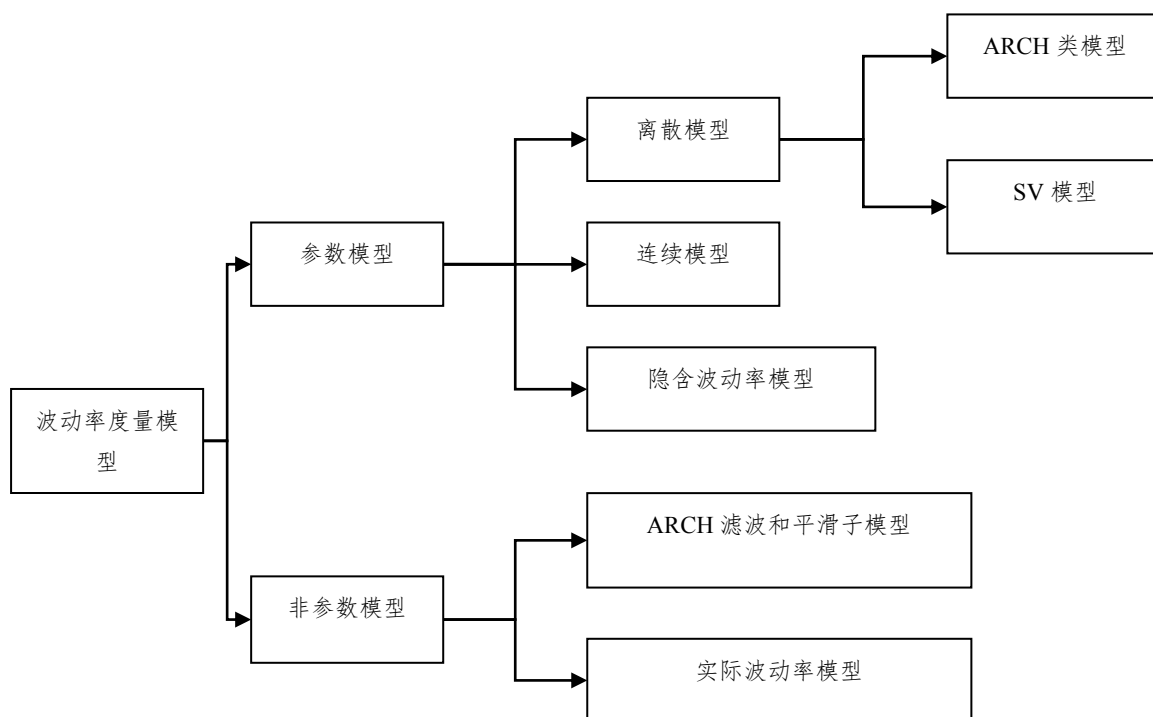
波动率的度量方法，可分为参数法和非参数法两类：参数法是利用一定的参数模型度量波动率，波动率变量是内嵌于模型内，如 ARMA、GARCH 和 SV 模型；非参数法

是利用日交易数据按一定的方法直接计算而得，如收益方差、日收益绝对值等。

近年来，度量波动率转向了利用高频率数据的非参数方法上。Andersen 等 (1998,2001) 提出了一种度量波动率的新方法，称之为实际波动率 (Realized Volatility)，是通过加总某一频率下的日内分时数据的收益平方来得到真实波动率的一个估计。

理论证明：在日内频率选取适当的情形下，该估计量是真实波动率的无偏一致且有效的估计量。因此，大量文献致力于利用高频样本数据来研究非参数的实际波动率。而对于最优样本频率的选取，则成为计算实际波动率最为关键的问题。若频率过小，则不会得到真实波动率一致的估计量；若频率过大，由于收益受到市场微观结构噪声的影响，度量结果会有较大误差。

ABDL (2001b) 提出了 VAR—RV 模型，即所谓的长记忆高斯向量自回归对数实际波动率模型，研究发现 VAR-RV 的预测精度优于 GARCH (1,1) 的预测精度。因为 GARCH (1,1) 用到的是直到 T-1 日的日收益平方，而 VAR-RV 利用的却是直到 T-1 日的日内收益数据，是基于长记忆的动态模型。



图表 1 波动率度量方法

## 二、实际波动率与 VIX 指数区别

### (一) VIX 指数

VIX 波动率指数是由 CBOE (芝加哥期权交易所) 在 1993 年所推出，是指数期权隐含波动率加权平均后所得的指数。最初，VIX 选取 S&P100 指数期权的近月与次月最接近平价的看涨期权及看跌期权共八个序列，分别计算其隐含波动率之后再加权平均所得。2003 年，VIX 编制方法转为方差互换方法进行计算。

VIX 反映了期权投资者对未来市场波动的预期，当指数越高时，显示投资者预期未来市场波动越剧烈；当 VIX 指数越低时，表示投资者认为未来市场波动将趋于缓和。由于透过 VIX 该指数可以观察期权参与者的心理，也被称为“投资者情绪指标”(The investor fear gauge)。经过十多年的发展和完善，VIX 指数逐渐得到市场认同，

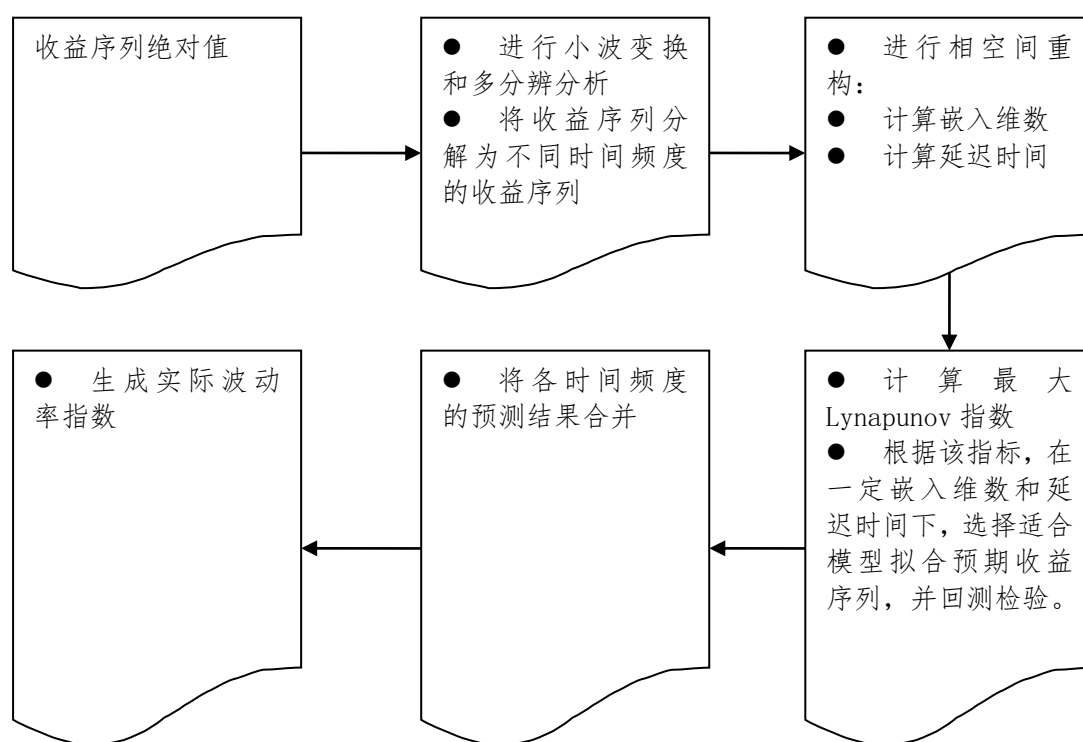
CBOE 于 2004 年推出了第一个波动率期货 (Volatility Index Futures) VIX Futures，2004 年推出第二个将波动率商品化的期货，即方差期货(Variance Futures)，标的为三个月期的 S&P500 指数的实际方差(Realized Variance)。2006 年，VIX 指数的期权开始在芝加哥期权交易所开始交易。

## (二) 实际波动率指数

实际波动率指数的理论来自于 Andrew Clark,et. (2009) 的研究，并且该指数获得了 2009 年指数超级杯颁发的重要奖项，

同 VIX 指数相比，实际波动率指数给出了一种对于实际波动率的度量方式，可以用来预测多时期范围的波动率，比如短期或者长期的实际波动率。

## 三、实际波动率指数计算



图表 2 实际波动率指数计算流程

### (一) 波动率序列

实际波动率指数采用收益绝对值作为波动率的表示。收益绝对值也是波动率度量的

一种重要表示方式。

## (二) 小波分解

传统预测方法是建立线性模型来模拟对象的发展规律，比较常用的模型有 AR 模型和 ARMA 模型，但前提是假设时间序列是平稳的。在现实中，时间序列往往是非平稳的，小波多分辨分析可以把时间序列分解为若干层，由于分解后的序列在频率成分上比源序列单一，并且对序列进行了平滑（去噪处理），因此平稳性比原序列要好。

实际波动率指数，首先就是对历史波动性序列进行小波分解成多期形式，并对分解后的单支序列通过相空间重构方法，确定嵌入维数和时间滞后项，最终采用传统方法对波动率进行预测，最后综合每层的预测值可得到原波动率序列的预测值，从而实现对实际波动率的预测。

### 1、离散小波变换和多分辨分析

设  $X$  为  $N$  维向量，元素为实值时间序列  $(X_t, t=0, \dots, N-1)$ ，样本数  $N$  是  $2^{M_0}$  的整数倍（ $M_0$  是进行小波变换的层数）。令  $\{W_n, n=0, \dots, N-1\}$  代表小波变换的系数，其表达式为：

$$W = \omega X$$

其中， $W = [W_1, W_2, \dots, W_{M_0}, V_{M_0}]^T$ ， $\omega = [\omega_1, \omega_2, \dots, \omega_{M_0}, v_{M_0}]^T$ 。

$W_m = \omega_m X, V_{M_0} = v_{M_0} X$ 。 $W_m$  是一个维数  $N_m = N/2^m$  的小波系数向量，随尺度  $\tau_m = 2^{m-1}$  的变化而变化。 $\omega_m$  是一个  $N_m \times N$  维矩阵， $V_{M_0}$  是一个维数  $N_{M_0}$  的尺度系数矩阵。它与尺度  $\lambda_{M_0} = 2^{M_0}$  有关。 $v_{M_0}$  是一个  $N_{M_0} \times N$  维矩阵。 $W_m$  和  $V_{M_0}$  是标准正交的， $\langle W_m, V_{M_0} \rangle = 0, m=1; \dots; M_0$ ， $M_0$  是进行小波变换的层数。

向量  $X$  可以由  $W$  来合成：

$$X = \omega^T W = \sum_{m=1}^{M_0} \omega_m^T W_m + v_{M_0}^T V_{M_0} = \sum_{m=1}^{M_0} D_m + S_{M_0}$$

将  $X$  分解为  $M_0+1$  个  $N$  维向量，其中， $D_m = \omega_m^T W_m$  ( $m=1, \dots, M_0$ )，简称  $D_m$  为对应于尺度  $\tau_m$  的  $m$  级细节， $S_{M_0} = v_{M_0}^T V_{M_0}$  为对应于尺度  $\lambda_{M_0}$  的  $M_0$  级平滑。

设  $X$  经过离散小波变换 (DWT) 变换后，变为  $M_0$  层， $W_m$  之间是相互正交的， $\langle W_m, V_{M_0} \rangle = 0 (m \neq n)$ ，则  $X$  被分解为：

$$\|X\|^2 = \sum_{m=1}^{M_0} \|W_m\|^2 + \|V_{M_0}\|^2$$

其中， $\|\bullet\|$  表示向量的模。

### 2. 最大重复离散小波变换 (MODWT)

最大重复离散小波变换可看作经过修正后的一种 DWT 形式。与 DWT 是正交变换不同，MODWT 是一个高度冗余的非正交变换。而且  $M_0$  级 MODWT 可以为任意样本容

量  $N$ ，而  $M_0$  级 DWT 的样本容量应满足  $N = C \times 2^{M_0}$  ( $C$  为任意常数)。因此，MODWT 比 DWT 的应用范围更广泛。

与离散小波转换的定义相同，向量  $X$  经过 MODWT 后，其表达式为：

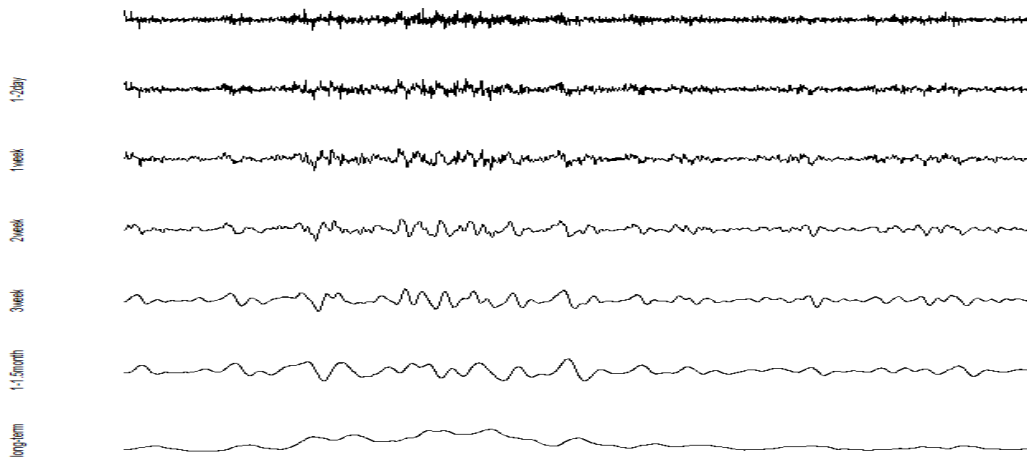
$$X = \omega^T W = \sum_{m=1}^{M_0} \omega_m^T W_m + v_{M_0}^T V_{M_0} = \sum_{m=1}^{M_0} D_m + S_{M_0}$$

其中， $W = [W_1, W_2, \dots, W_{M_0}, V_{M_0}]^T$ ， $\omega = [\omega_1, \omega_2, \dots, \omega_{M_0}, v_{M_0}]$ 。

$W_j = [W_{j,0}, W_{j,1}, \dots, W_{j,N-1}]^T$  ( $j=1, 2, \dots, M_0$ )， $V_{M_0} = [V_{M_0,0}, V_{M_0,1}, \dots, V_{M_0,N-1}]^T$ 。

$W_m = \omega_m X$ ,  $V_{M_0} = v_{M_0} X$ 。  $D_m$  和  $S_{M_0}$  分别是经过最大重复离散小波变换后所得到的细节部分和平滑部分。

采用 MODWT 算法，选用 HAAR 小波基，沪深 300 波动率小波分解如下



图表 3 沪深 300 波动率分解

### (三) 非线性动力系统

通常，单变量时间序列仅能提供系统有限的信息，实际波动率时间序列原则上是复杂的动态系统中各要素相互作用的结果。Packard 和 Takens 等提出选取合适的嵌入维数  $m$  和时间延迟  $\tau$ ， $X(t_i) = [x(t_i), x(t_i + \tau), x(t_i + 2\tau), \dots, x(t_i + (m-1)\tau)]$  ( $i=1, 2, \dots, N$ )，即通过对一维时序的维数进行扩充和延拓，将其扩展到三维甚至更高维的空间中去，将时间序列中蕴藏的信息充分地显露出来。这样可在拓扑等价的意义下恢复原吸引子的动力学特征，就能由系统的当前状态获得下一时刻的状态，从而得到时间序列下一时刻的预测值，这为时间序列的预测提供了依据。重构相空间的关键在于嵌入维数  $m$  和时间延迟  $\tau$  选取。

#### 1、延迟时间选取

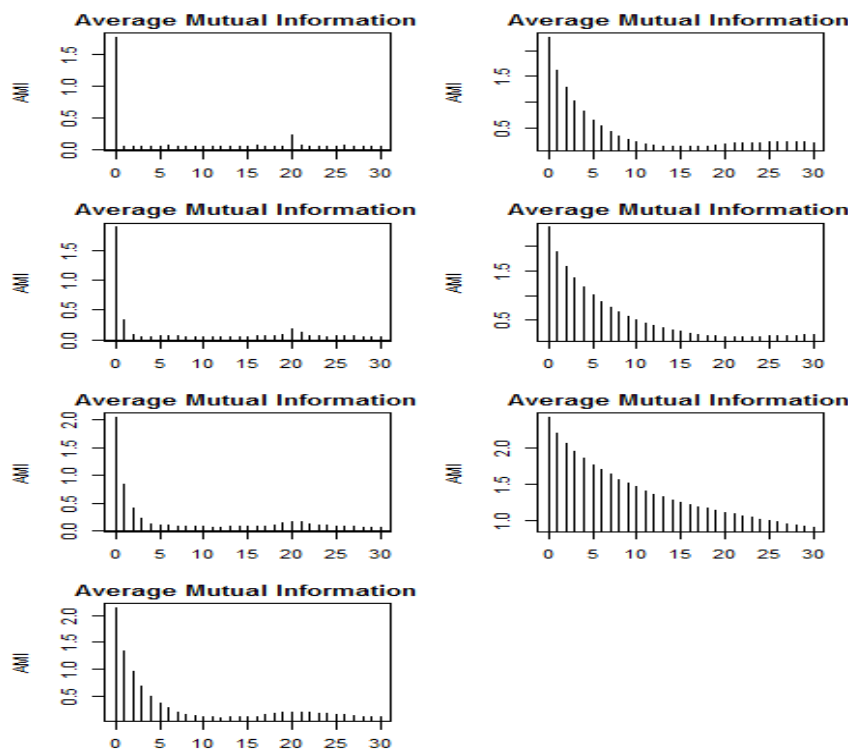
延迟时间  $\tau$  是的选择方法主要有自相关函数法、重构展开法、最小互信息法等。通常采用最小互信息法。

最小互信息法是构造互信息函数，其中包含了时间序列的非线性特征，从事件  $b_j$  在序列  $B$  中发生而得到的关于事件  $a$  在序列  $A$  中发生的信息，称为  $A$  和  $B$  之间的平均互信息，关系如下：

$$I(A, B) = \frac{1}{N} \sum_{i,j} P_{AB}(a_i, b_j) \times \log \left[ \frac{P_{AB}(a_i, b_j)}{P_A(a_i)P_B(b_j)} \right]$$

在实际计算时通常采用划分网格的方法，将变量和组成的样本空间划分为若干网格或盒子，然后通过统计各盒子中的点数来求出其概率值，一般选取延迟矢量之间的互信息函数第 1 个极小值点的时间为优化的延迟时间。

通过最小互信息法，确定沪深 300 小波分解后每个波的滞后阶数为 9, 8, 11, 12, 14, 22, 30。



图表 4 延迟时间确定

## 2、最佳嵌入维数

嵌入维数是相空间重构中又一重要参数。如果嵌入维数较小，吸引子将折叠，某些原本离得较远的点可能离得较近；如果嵌入维数太大，需要更多的数据，且噪声可能占主导地位。因此，最佳嵌入维数的选取非常重要。目前，嵌入维数选取方法主要有 3 种：饱和关联维数法、伪近邻点法及奇异值分解法等。通常采用伪近邻点法(FNN)计算嵌入维数。

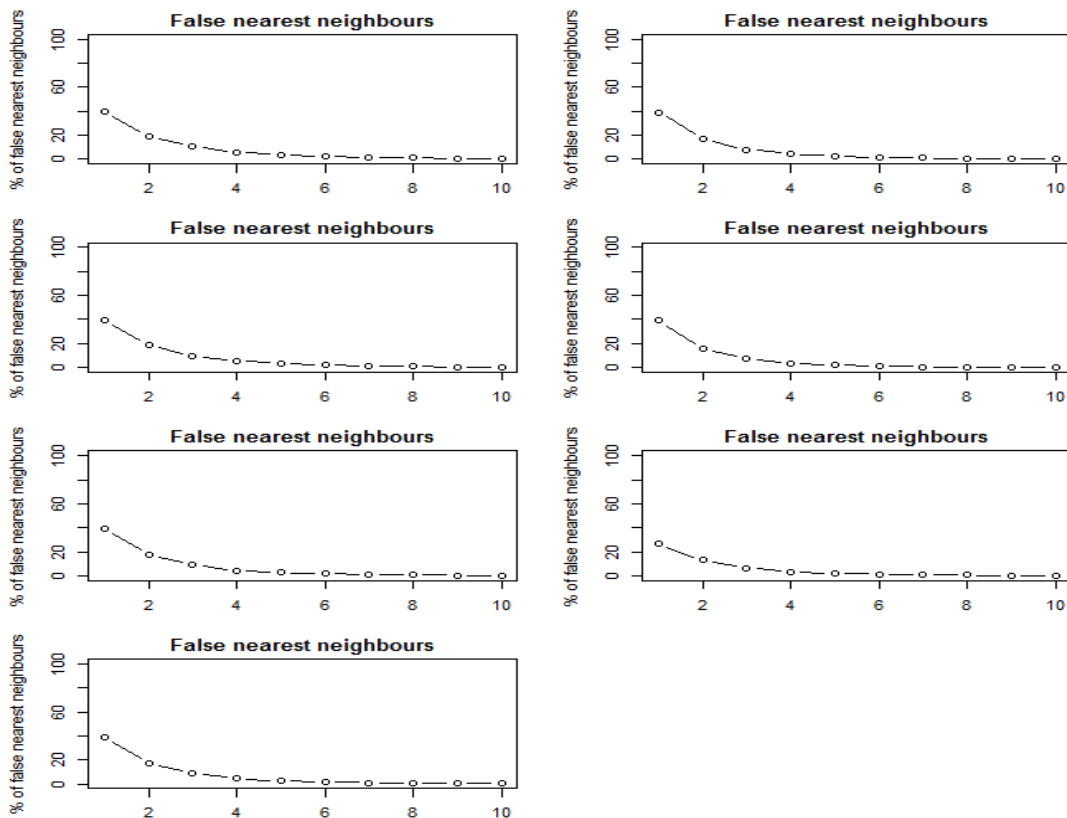
伪近邻点法是一种基于吸引子空间几何结构是否被完全打开的方法。嵌入维数选的较小相当于吸引子投影到低维空间，将使得原本离得较远的点成为近邻点，因而产生“伪近邻点”，可以通过判断伪近邻点的比率来选取合适的嵌入维数。设重构延迟矢量  $X_j$  的



近邻点为  $X_j$ ，根据欧式空间理论， $X_i$  和  $X_j$  之间的距离为  $R_i(m) = \|X_i - X_j\|^{(m)}$ ，维数增大一维到  $m+1$  时距离为  $R_i(m+1)$ ，如果  $R_i(m+1)$  比  $R_i(m)$  大很多，则可推断  $X_i$  和  $X_j$  是伪近邻点。在计算时，根据实际情况选择临界值  $r$ ，看是否满足下列不等式

$$\frac{R_i(m+1) - R_i(m)}{R_i(m)} > r$$

由此来判断  $X_i$  和  $X_j$  是否为伪近邻点，从  $m=2$  开始取值，取一固定阈值，计算虚假最近邻点的比例，然后增加  $m$ ，直到虚假最近邻点小于或者最近邻点数不再随  $m$  的增加而减少时，此时的  $m$  为最小嵌入维数。



图表 5 最佳嵌入维数确定

#### (四) 最大 Lyapunov 指数计算

在确定了嵌入维数和延迟时间两个重要参数之后，就需要计算最大李雅普诺夫指数 ( $\lambda_{\max}$ )。李雅普诺夫指数决定了波动率预测模型的选择。

如果  $\lambda_{\max} < 0$ ，则表示相邻点最终要合并为一点，这对应于稳定的不动点，因此，有很多线性和非线性的模型可以选取来进行波动率预测；

如果  $\lambda_{\max} = 0$  或者接近于 0，则对应周期运动，可以选择周期或者近周期函数进行预测；

如果  $\lambda_{\max} > 0$ ，则意味着相邻点最终要分离，对应着轨道的局部不稳定，表示系统存

在混沌行为，可采用混沌序列预测方法。

### 1、最大李雅普诺夫指数计算

李雅普诺夫指数 (Lyapunov Exponents)  $\lambda$  可以表征系统运动的特征，它沿某一方向取值的正负和大小，表示长时间系统在吸引子中相邻轨道沿该方向平均发散 ( $\lambda_i > 0$ ) 或收敛 ( $\lambda_i < 0$ ) 的快慢程度。因此，最大的李雅普诺夫指数  $\lambda_{\max}$  决定轨道覆盖整个吸引子的快慢，最小李雅普诺夫指数  $\lambda_{\min}$  则决定轨道收敛的快慢。

对混沌动力系统， $\lambda$  的大小与系统的混沌程度有关，假设系统从相空间中某半径足够小的超球开始演变，则第  $i$  个李雅普诺夫指数定义为：

$$\lambda_i = \lim_{t \rightarrow \infty} \log(r_i(t)/r_i(0))$$

其中， $r_i(t)$  为  $t$  时刻按长度排在第  $i$  位的椭圆轴的长度； $r_i(0)$  初始球的半径。

计算最大李雅普诺夫指数的主要步骤为：

- (1) 应用实际时间序列重构一个高维的相空间及相空间中的点集；
- (2) 以初始相点  $x_0$  为基点，在重构的相空间中选取一个与  $X_0$  相距至少一个轨道周期的最近点作为端点，构成一个初始向量  $V_0^\mu$ ，求出该向量的长度，记为  $L_0$ ；
- (3) 经过一个进化时间  $\tau_0$ ，初始向量  $V_0^\mu$  运动发展为另一个向量  $V_1^\mu$ ，其相应的起点和端点分别为  $x(t_0 + \tau_0)$  和  $y(t_0 + \tau_0)$ ，计算出其相长度为  $L_1$ ，相长度在时间  $\tau_0$  内由  $L_0$  变为  $L_1$ ，若用  $\lambda_1$  表示在此时间段内相长度的指数增长率，则有：

$$L_1 = L_0 e^{\lambda_1 \tau_0}$$

- (4) 以  $x(t_0 + \tau_0)$  为新的基点，选取一个新的向量  $V_1^\mu$  为发展向量，发展向量具有小的长度并与  $V_1^\mu$  保持较小的夹角，再以  $V_1^\mu$  为初始向量，用同样的方法，可得到指数增长率  $\lambda_2$ ：

$$\lambda_2 = \frac{1}{\tau_1} \log(L_2/L_1)$$

上述过程一直发展到点集终点，取增长率  $\lambda_i$  的平均值作为最大李雅普诺夫指数的估计值，即：

$$\lambda = \frac{1}{N_k} \sum_{k=1}^{N_k} \frac{1}{T_{k-1}} \ln(L_k/L_{k-1})$$

- (5) 增减嵌入维数  $m$ ，重复上面计算步骤，直到指数估计随  $m$  保持平衡为止，最终所得计算结果即为所要求的最大李雅普诺夫指数。

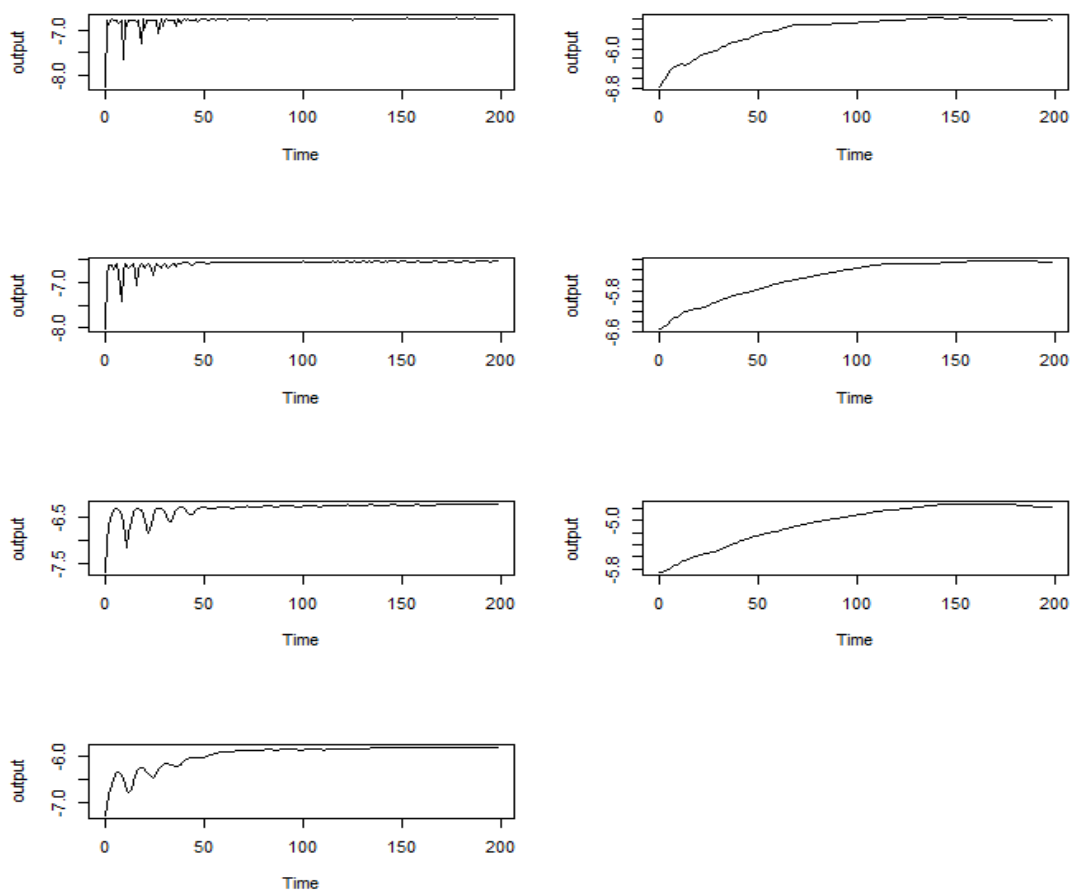
### 2、混沌序列预测

用相空间重构理论来预测时间序列有多种方法，从拟合相空间中吸引子的方式来看，混沌预测可分为全域预测法和局域预测法两种。

全域预测法是将轨迹中的全部点作为拟合对象，找出其规律，由此预测轨迹的走向。

局域预测法是在全域预测法的基础上改进而来的。

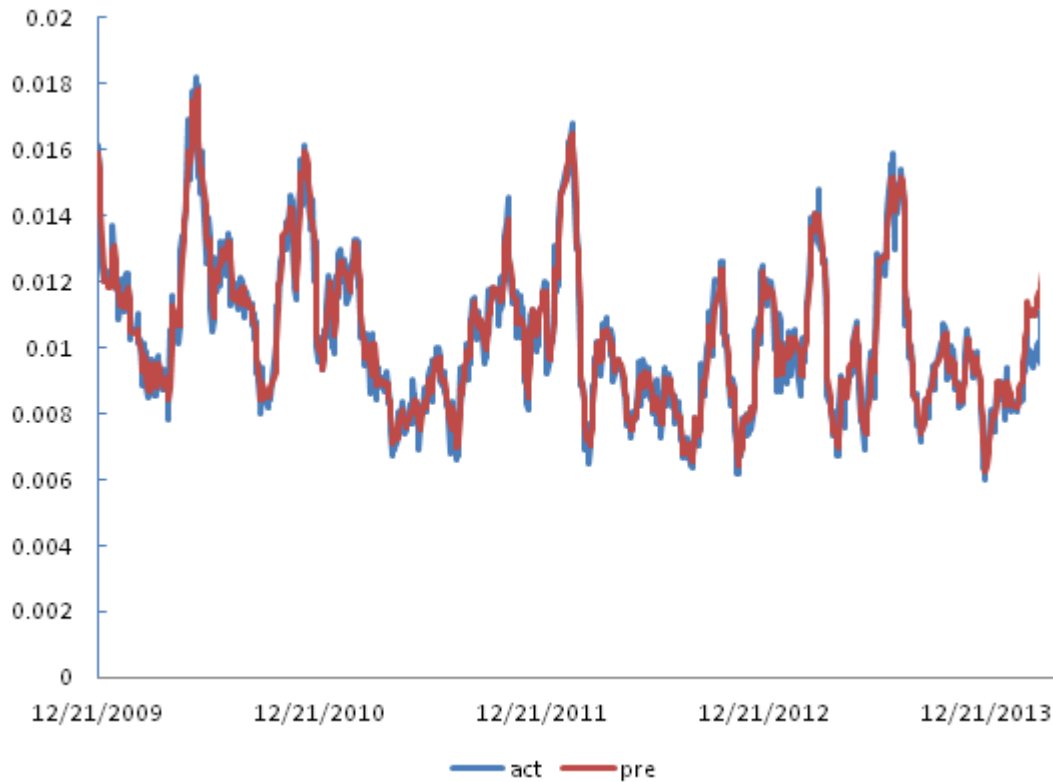
混沌局域预测法则是从全部样本点中选取那些离中心点最近的轨迹点作为相关点做出拟合。这些相关点的选取对混沌预测模型的建立以及预测结果是有影响的。因此，选取哪些轨迹点作为相关点是一个问题。在计算各点到中心点的距离时，可以有多种方法，如欧式距离法（Euclidean Distance），余弦距离法（Cosine Distance），而对数据点进行拟合时，可以采用局域线性预测模型，局域常数预测模型，局域加权线性预测模型等。



图表 6 最大李雅普诺夫指数计算

### （五）实际波动率指数生成

通过对波动率序列进行小波分解后，分解出不同时间层级的平稳序列，然后针对每层的序列进行重构和拟合，并进行预测。最后，综合每层的预测值可得到原波动率序列的预测值。从而生成实际波动率指数。



图表 7 实际波动率指数表现

#### 四、结论

实际波动率指数给出了一种不同于 VIX 的波动率指数。该指数所运用的小波分解理论和动态系统理论，国内外学者都进行了大量的研究，而且很多学者采用类似的方法进行波动率预测，但结果都不大理想。

此次运用上述方法编制的波动率指数从已有的资料来看，效果不错。不过实际波动率指数，对于波动率的收益绝对值的表示形式，小波分解采用的最大重复离散小波变换方法，确定嵌入维数的伪近邻点方法，确定延迟时间的最大互信息法，进行混沌序列预测采用的局域预测法，这些方法都对最终的波动率预测起着至关重要的作用。因此，对于实际波动率指数的最终适用性还需要进一步的检验。