

# 霍克斯过程及其在高频建模中的应用

Baron Law and Frederi G. Viens

*Purdue University, West Lafayette, IN, USA*

## 7.1 介绍

---

在这一章中介绍并概述了一系列新兴的随机点过程（**stochastic point process**）。这方法正被用于给股票市场上大量的量化高频数据建模上。

股票市场的限价指令簿（**LOB**）包含的信息是一个多元的时间序列。这个时间序列记录了股票市场中数千只股票的订单发起时间和每个价格水平下的需求量。限价指令簿有着一些明显特征[1-3]，包括

1. 订单发起的不规则时间间隔
2. 价格变化和订单量变化的离散状态空间
3. 一日内的季节性变化（开市和闭市时有更多的交易活动）
4. 订单发起的群聚
5. 历史的自激效应
6. 来自其他资产的交叉激励效应
7. 激励效应的持久性

因此，具有固定时间间隔的时间序列模型，例如 **ARIMA** 和 **GARCH**，不能很好的用于高频模型建模。一个在实践中广泛应用的方法是把数据以 5 分钟为间隔重新采样[4,5]，因此解决了流动性强的股票的

时间测度，这样许多前述的特征都可以观察到。但这可能总共舍弃了这种股票 99% 的数据。相反，泊松过程（Poisson processes），这种广泛用于市场微观结构的方法[6,7]不能很好的刻画前述高频数据的普遍特征。

这个概述论文，概述目前在高频金融数据建模的研究，关注霍克斯过程（Hawkes process），点过程的一种，用于有自激效应和交叉激励效应的建模。在 7.2 节，我们提供了一个非正式的对于点过程的介绍。更专业的材料可以在附录 7. A 中找到。所有关于点过程的内容大多都包含在主流教材中[8-12]。7.3 和 7.4 节介绍了霍克斯过程和他的统计学结论，附录 7. B 提供关于他的历史的一个简短的概述。7.5 小节展示了霍克斯过程在高频数据建模上的应用。

## 7.2 点过程

---

这节提供了一个对于点过程的非正式介绍。一个用随机测度(random measures)和鞅论(martingale theory)的更严格的处理可以在附录 7. A 中找到。

一个点过程是在集合  $X$  中的一个随机的可数点集  $\{x_1, x_2, \dots\}$ 。  $\{x_n\}$ 's 是  $X$ -valued 随机变量。他们可以有高度不平凡的相关性结构。比如说  $\{x_n\}$ 's 可以是  $X = \mathbb{R}^2$  空间中的地震坐标，并且如果自激性质被引入到这个模型中，这就是我们要讨论的霍克斯过程，  $\{x_n\}$ 's 可以表示余震的随机分布。

$N(A)$ 代表在区域 $A \subseteq X$ 中的点的数量。如果对于  $X$  的任何子集  $A$ ，都有 $N(A)$ 是知道的，那么这在基本上的决定了所有点的位置，并且这个映射是用来描述点过程的最基本工具。只有当每个坐标至多含有一个点，点过程  $N$ 才被称为简单 (simple)。

当  $X$  是数轴正半轴 $R_+$ 时，典型情况下，这代表时间轴。点被看做事件发生的时间。 $N(t) = N((0, t])$ 表示了包括或在时间点  $t$  之前发生事件的次数。由于我们的目标是给高频数据建模，我们将主要应对在 $R_+$ 上的点过程。

当一个事件发生在 $t_n \in R_+$ 时，这可能带有一个额外的信息  $y_n$ （它的标记）。在金融中，每一个订单的产生都伴随着它的订单量。在地质学中，每个地震的发生都伴随着它的震级。这种伴随着标记的点过程被称为标记点过程 (marked point process, MPP)。让  $Y$  代表标记空间，并且 $A \subseteq Y$ ，那么

$N(t, A) = N((0, t] \times A)$ 代表了发生在时间  $t$  及之前，并且标记落在了区域  $A$  中的点的数目。Ground process  $N_g(t) =$

$N((0, t] \times Y)$ 表示了发生在 $(0, t]$ 的事件总次数，无论他们的标记是什么。当 $y_n$ 独立于 $\{(t_i, y_i)\}_{i < n}$ 时，这个 MPP 的标记被称为不可预测。当 $y_n$ 独立于 $\{(t_i, y_i)\}_{i \neq n}^2$ 时，MPP 的标记被称为独立的。当一个 MPP 的 ground process  $N_g$ 是简单的，MPP 被称为简单。

如果一个点过程在某个坐标处出现了多个点，我们可以把他当作带有点数目标记的简单 MPP。因此，不失一般性的，大部分的情况都将基于简单点过程。

当标记空间 $Y = \{1, \dots, d\}$ 时， $N$ 被称为多变量点过程。在这种情况下， $N_i(\bullet) = N(\bullet \times \{i\})$ 被称为边际点过程。一个  $n$  变量

的简单点过程和一个  $n$  维简单点过程不同。前者不能有 **common jump times**。如果一个多变量的在每一点有额外的信息  $w_n \in W$ 。标记空间将成为  $Y = \{1, \dots, d\} \times W$ 。边际过程  $N_i(t, A) = N((0, t] \times \{i\} \times A)$ 。边际 ground process 是  $N_i(t) = N((0, t] \times \{i\} \times Y)$ 。

许多点过程可以以他们的随机强度 (stochastic intensities)  $\lambda(t)$ , 这被不严格的定义为考虑  $t$  时刻前的所有信息得出的  $t$  时刻每单位时间内订单的预期到达数量。表达式为:

$$\lambda(t) = \lim_{h \rightarrow 0^+} \frac{E(N((t-h, t]) | F_{t-h})}{h}$$

而  $F_t$  代表所有包含在到时间  $t$  为止的信息。对于一个多变量的点过程,  $\lambda_i(t)$  是边际点过程  $N_i(t)$  的强度。

最有名的点过程是泊松过程, 他的强度是由时间决定的函数。当一个过程产生强度是一个随机过程时, 他就和扩散中的随机波动很相似。一个随机强度的重要用途是允许它可以根据事件的演变而改变。霍克斯过程拥有这样一个特征, 以完成含有自激效应的建模。

### 7.3 霍克斯过程

---

霍克斯过程是一种随机强度有着自迭代形式的点过程。对于一个非线性的多元标记霍克斯过程, 他的强度  $\lambda(t) = (\lambda_1(t), \dots, \lambda_d(t))$  由如下给出

$$\begin{aligned}
\lambda_i(t) &= \Phi_i \left( \sum_{j=1}^d \int_{(-\infty, t) \times Y} \gamma_{ij}(t-s, y) N_j(ds \times dy), t \right) \\
&= \Phi_i \left( \sum_{t_n < t} \gamma_{i, w_n}(t-t_n, y_n), t \right) \\
\Phi_i &: \mathbb{R} \times \mathbb{R}_+ \longrightarrow \mathbb{R}_+, \quad \gamma_{ij}: \mathbb{R}_+ \times Y \longrightarrow \mathbb{R}, \\
N_j &: \mathcal{B}(\mathbb{R}_+ \times Y) \longrightarrow \mathbb{N}
\end{aligned} \tag{7.2}$$

其中  $w_n \in \{1, \dots, d\}$  标识  $t_n$  的类型  $\Phi_i$  是比率函数。考虑一个特殊的情况

$$\begin{aligned}
\lambda_i(t) &= \mu_i(t) + \sum_{j=1}^d \int_{(-\infty, t) \times Y} \gamma_{ij}(t-s, y) N_j(ds \times dy) = \mu_i(t) \\
&\quad + \sum_{t_n < t} \gamma_{i, w_n}(t-t_n, y_n) \\
\mu_i &: \mathbb{R}_+ \longrightarrow \mathbb{R}_+, \quad \gamma_{ij}: \mathbb{R}_+ \times Y \longrightarrow \mathbb{R}_+, \\
N_j &: \mathcal{B}(\mathbb{R}_+ \times Y) \longrightarrow \mathbb{N}
\end{aligned} \tag{7.3}$$

这是  $\Phi_i(x, t) = \mu_i(t) + x$ 。这种由 (7.3) 确定的霍克斯过程为线性的。 $\mu_i(t)$  被称为基比率。函数  $\gamma_{ij}$  被称为衰变核 (decay kernel)，通常  $\gamma_{ij}(t, y)$  有着分散的形式  $\gamma_{ij}(t)g_{ij}(y)$ ，其中  $g_{ij}$  被称为影响核。通常衰变核的选择有  $\alpha_{ij}e^{-\beta_{ij}t}$ ， $\alpha_{ij}(c_{ij} + t)^{-\beta_{ij}}$ ， $\sum_{k=0}^K \alpha_{ijk}t^k e^{-\beta_{ijk}t}$ 。

如果衰变函数有着指数形式，同时  $\beta_{ij} = \beta_i$ ，强度  $\lambda(t)$  与矢量  $(N(t), \lambda(t))$  都是马尔科夫过程，并且保证  $\mu_i(t) = \mu_i$ ，那么  $(Y^1(t), \dots, Y^d(t))$  满足随机微分方程等式

$$d\lambda_i(t) = \beta_i(\mu_i - \lambda_i(t))dt + \sum_{j=1}^d \alpha_{ij}dN_j(t) \tag{7.4}$$

这个规范有着一个简单的演绎。发生在时间  $t$  前的事件  $N_j$  通过  $\alpha_{ij} \geq 0$ ，使得强度增加这样触发了未来的事件。如果强度  $\lambda_i(t)$  比  $\mu_i$  要高，第一项变为了负的，使得强度被拉回平衡水平  $\mu_i$ 。换句话说，

强度  $\lambda_i(t)$  是一个在均值附近往复的过程，它由它的点过程驱动。马尔科夫属性和这个直观的演绎或许可以解释为什么指数的衰退核会被如此广泛的使用。

对于线性霍克斯过程， $\lambda_j(t)$  的， $\lambda_{ij}(t)$  和  $g_{ij}$  必须为正数对于所有的  $t$ ，以确保  $\lambda(t)$  为正。其结果是，与非线性霍克斯过程，线性霍克斯过程不能建模的抑制效果（负激励）。然而，线性霍克斯过程更容易实现。它们的性质被更好的理解，而最重要的是，它们具有分支结构的表示，这是在模拟，估计和模型的解释是非常有用的。

---

<sup>6</sup> $N$  本身并不是一个马尔可夫过程作为其在时间  $t$  的强度取决于它的全部历史前时间  $t$ 。

### 7.3.1 分支结构表征

线性霍克斯过程有一个非常优雅的分枝结构表征。我们在这里描述的版本为不可预测的标记多元霍克斯过程，

有  $d$  种移民，根据泊松过程，他们有着比率  $\lambda_1, \dots, \lambda_d$  出生或抵当每个人（或后裔移民），将携带一个不可预知的标志。在时间  $t_n$  出生的  $j$  型带有标志  $y_n$  的个体会生出  $i$  型的个体

根据带有比率  $\gamma_{ij}(t - t_n, y_n)$  的非齐次泊松过程（T-TN, YN）。所有非齐次泊松过程是相互独立的。所有的非齐次泊松过程是相互独立的。

令  $N_i(t)$  为  $i$  型诞生的个体总数/到达时间  $t$  或之前出现上述情况下，那么  $N(t) = (N_1(t), \dots, N_d(t))$  将遵循线性标注霍克斯过程(7.3) 这种表现形式在 7.4.2 节期望最大化 (EM) 算法的基础上，我们也将看到它是如何用于测量第 7.5.4 节一分过程的内生性。

### 7.3.2 平稳

考虑一个霍克斯过程  $N$  带有强度 (7.2) 使得  $\Phi_i(x, t) = \Phi_i(x)$ ,  $N$  有一个特殊的平稳版本，如果如下的任意一条件被满足 [18, 20]:

1.  $\Phi_i(x)$  是  $k_i$ -Lipschitz 并且谱半径  $\lambda(A) < 1$  对于  $d \times d$  阶矩阵

$$A = [k_i \int_0^\infty |\gamma_{ij}(t)| dt]_{i,j}.$$

2.  $\Phi_i(x)$  是 Lipschitz,  $\Phi_i(x) \leq M$ ,  $\int_0^\infty |\gamma_{ij}(t)| dt < \infty$  并且

$$\int_0^\infty t |\gamma_{ij}(t)| dt < \infty.$$

从技术上来说， $N$  可具有其它非固定的版本与固定一起;然而，非平稳版本将弱收敛到固定的版本时  $t \rightarrow \infty$ . 因为霍克斯过程从  $-\infty$  开

始,对于所有的  $t > 0$   $N((0, t])$  会有着平稳分布.

### 7.3.3 收敛性

这一节中,我们将讲述多种霍克斯过程的收敛性的结果。如果一个强档的有标度的线性霍克斯过程衰变函数的谱半径小于 1, 它将弱收敛于伯努利扩散。当它的谱半径接近于 1, 他将收敛在一个 Cox–Ingersoll–Ross (CIR) 过程。对于非线性的霍克斯过程, 我们只有根据 Lipschitz 常数  $\Phi$  所得的单变量情况下根据的结果和充分条件。

#### 7.3.3.1 对于多变量霍克斯过程的大数定律

假设一个没有标记的模型, 如果它的谱半径  $\rho(A) < 1$ , 其中

$A = [\int_0^\infty \gamma_{ij}(t)dt]_{i,j}$ , then [24], 那么

$$\sup_{t \in [0,1]} \left\| \frac{N(nt)}{n} - t(I_d - A)^{-1} \mu \right\| \xrightarrow[n \rightarrow \infty]{\text{a.s./}L^2} 0^{10,11} \quad (7)$$

其中  $\mu = (\mu_1, \dots, \mu_d)$ . 当  $d = 1$  并且我们取  $t = 1$ , 这表明了

$$\frac{N(T)}{T} \xrightarrow[T \rightarrow \infty]{\text{a.s./}L^2} \frac{\mu}{1 - \int_0^\infty \gamma(t)dt} \quad (7.6)$$

#### 7.3.3.2 对于多变量霍克斯过程的中心极限定理

假设一个没有标记的模型,  $N = (N_1, \dots, N_d)$ , 如果谱半径  $\rho(A) < 1$ , 其中

$$A = [\int_0^\infty \gamma_{ij}(t)dt]_{i,j} \text{ and } \int_0^\infty \sqrt{t} \gamma_{ij}(t)dt < \infty \quad \forall i, j,$$

, 那么

$$\sqrt{n} (N(\bullet n)/n - \bullet(I_d - A)^{-1} \mu) \xrightarrow[n \rightarrow \infty]{\text{weak}} (I_d - A)^{-1} \Sigma^{1/2} W(\bullet)^{12} \quad (7)$$

$\Sigma = \text{diag}((I_d - A)^{-1} \mu)^{13}$ ,  $W$  是一个标准  $d$  维的伯努利运



动。

### 7.3.3.3 对于单变量非线性霍克斯过程的中心极限定理

假设一个没有标记的模型(7.2) 并且  $d = 1$ , 如果  $\gamma(t)$  在下降,

$$\int_0^\infty t\gamma(t)dt < \infty, \quad \Phi(x,t) = \Phi(x) \text{ 在上升, 并且是 } k\text{-Lipschitz}$$

$$\int_0^\infty k\gamma(t)dt < 1, \quad \text{那么,}$$

$$\sqrt{n}(N(\bullet n)/n - \bullet v) \xrightarrow[n \rightarrow \infty]{\text{weak}} \sigma W(\bullet) \quad (7.9)$$

$$\sigma^2 = \mathbb{E}((N([0, 1]) - v)^2) + 2 \sum_{n=1}^{\infty} \mathbb{E}((N([0, 1]) - v) \times (N([n, n+1]) - v), \quad v = \mathbb{E}(N([0, 1])) \quad (7.10)$$

## 7.4 霍克斯过程的统计学推论

### 7.4.1 仿真

这节中, 我们将给出一个仿真霍克斯过程的算法, 假设我们已知  $\mu(t)$  和  $\gamma(t, y)$  函数里所有的参数。我们的目标是仿真所有在区间  $[0, T]$  内的点  $(t_1, y_1), (t_2, y_2), \dots$ 。

如果标记分布只在  $t_n$  上, 我们可以简单的有条件的产生  $y_n$  在对应的  $t_n$  上, 然后  $t_{n+1}$  可以由强度函数  $\lambda(t)$ ,  $t > t_n$  产生, 这取决于  $\{(t_1, y_1), \dots, (t_n, y_n)\}$ . 如果  $y_n$  的分布也取决于  $\{(t_{n-1}, y_{n-1}), (t_{n-2}, y_{n-2}), \dots\}$ , 这个算法就可被适用。

#### 7.4.1.1 逆 CDF 变换

假设强度是被单变量霍克斯过程所决定的。让  $t_n$  代表到达时间,  $\tau_n = t_n - t_{n-1}$  是间隔时间。通过(7.63),  $\lambda(t) = h_n(t - t_{n-1})$  对于  $t \in (t_{n-1}, t_n]$ , 当  $h_n(t) = g_n(t) / (1 - G_n(t^-))$  且对于给定的  $\mathcal{F}_{t_{n-1}}$ ,  $g, G$  是  $\tau_n$  的条件 pdf, cdf。

如果  $G_n(t)$  是连续的,  $h_n(t)$  是简单的冒险函数 (hazard function), 那么有

$$G_n(\tau_n) = 1 - \exp\left(-\int_0^{\tau_n} h_n(s)ds\right) = 1 - \exp\left(-\int_{t_{n-1}}^{t_{n-1}+\tau_n} \lambda(s)ds\right) \quad (7.18)$$

给定的  
 $t_{n-1}$ , 我们  
 可以产生  
 $t_n = t_{n-1} +$   
 $\tau_n$  通过逆  
 cdf 变换

$\tau_n = G^{-1}(U)$ ,  $U \sim \text{Unif}(0, 1)$ . 然而, 这种倒置需要被数字上的表示, 所以这种方法被 Ogata 的修饰缩减 (Ogata's modified thinning) 所广泛替代。

#### 7.4.1.2 Ogata 修饰缩减 (modified thinning)

Ogata 介绍了修饰缩减方法, 这种方法不需要数字上的倒置。The 算法是基于以下的理论。让  $N = (N_1, \dots, N_d)$  成为一个有着强度  $(\lambda_1, \dots, \lambda_d)$  的多元点过程。

使得  $\sum_{i=1}^d \lambda_i(t) \leq \lambda^*(t) \forall t$  ( $\lambda^*(t)$  是外生的指定比率函数) 并且  $N^*$  是单一变量的带有强度  $\lambda^*(t)$  的非齐次泊松过程如果每个在  $N^*$  中的点  $t_n$  都带有标记  $y_n$  使得  $\mathbb{P}(y_n = i) = \lambda_i(t_n) / \lambda^*(t_n)$ ,  $i = 1, \dots, d$ , 那么  $(N^*, \dots, N^*)$  有着和  $(N_1, \dots, N_d)$  相同的分布

*Ogata's Modified Thinning*

1.  $n = 1$ ,  $t_0 = 0$
2. 产生:  $\tau_n \sim \text{Exp}(\lambda_n^*)$  for some  $\lambda_n^* \geq \sum_{i=1}^d (\lambda_i(t_{n-1}) + \alpha_i)$ .
3. 让  $t_n = t_{n-1} + \tau_n$ .
4. 产生  $U_n \sim \text{Unif}(0, 1)$ .
5. 如果  $U_n \in (\sum_{i=1}^{k-1} \lambda_i(t_n) / \lambda_n^*, \sum_{i=1}^k \lambda_i(t_n) / \lambda_n^*]$  for some  $k \in \{1, \dots, d\}$  返回  $t_n$  并且点有着  $k$  类型, 否则舍弃  $t_n$  (但保留值用来进行下一轮的步骤 2 和 3)
6.  $n = n + 1$ , 返回步骤 2

#### 7.4.1.2 分支结构仿真

这种方法使用线性标记霍克斯过程的分支结构表示生成点。j 型移民的到达时根据一个带有比率  $\mu_j(t)$  的非齐次泊松过程, 接下来, j 型的父辈在时间  $t_n$  到达, 带有标记  $y_n$ , 产生了 i 型的后代, 根据带有比率  $\gamma_{ij}(t - t_n, y_n)$  的非其次泊松过程, 产生重复每一个后代, 直到所有的人都超过预定时间  $T$ 。由于所有的非齐次泊松过程是独立的, 每一代可以并行完成。

1. 在  $[0, T]$  上产生带有强度  $\mu_i(t)$ ,  $i = 1, \dots, d$  的非齐次泊松过程。
2. 对于每个点  $t_n, y_n | t_n$
3. 假设是  $t_{nj}$  型的, 并且产生 i 型后代通过带有强度  $\gamma_{ij}(t - t_n, y_n)$  的非其次泊松过程
4. 重复第 2 步和第 3 步

The nonhomogeneous Poisson process with intensity  $(t)$  on  $[0, T]$  can be generated using Lewis' thinning algorithm

1. 产生  $N \sim \text{Poisson}(\mu^*)$  对于一些  $\mu^* \geq \max_{t \in (0, T]} \mu(t)$ .
2. 产生  $U_n \sim \text{Unif}(0, 1)$ ,  $n = 1, \dots, N$
3.  $T_n = U_{(n)}T$ ,  $n = 1, \dots, N$  ( $\{U_{(n)}\}$  是数据  $\{U_n\}$  的顺序)
4. 产生  $V_n \sim \text{Unif}(0, 1)$ ,  $n = 1, \dots, N$
5. 返回  $T_n$  if  $V_n \leq (\mu(T_n) / \mu^*)$ ,  $n = 1, \dots, N$ , 否则丢弃  $T_n$

#### 7.4.2 估计

假设我们观察一个点过程  $(0, T]$  和收集事件时间和标记  $\{(t_1, y_1), \dots, (t_N, y_N)\}$ , 现在我们

想在强度为  $\lambda(t)$ ，其驱动过程  $N(t)$  来估计函数  $\mu(t)$  和  $\gamma(T, Y)$ 。我们将总结出文献中的各种方法，但到目前为止，重点是在无人监视的进程。在其中标记是独立同分布的特殊情况下，该标记分布可以分别从点处理估计。

如果我们假设  $\mu(t)$  和  $\gamma(t)$  有一定的参数表示，我们可以用最大似然估计（MLE），EM，或者广义矩估计（GMM）来估计参数。否则，我们需要依靠一些先进的非参数技术来估计整个函数曲线

#### 7.4.2.2 最大似然估计

一个霍克斯过程的对数似然由下式给出

$$\log(L(\theta)) = \sum_{i=1}^d \left( - \int_0^T \lambda_i(t; \theta) dt + \int_0^T \log(\lambda_i(t; \theta)) dN_i(t) \right) \quad (7.19)$$

在多元线性霍克斯处理的情况下，就成了

$$\begin{aligned} \log(L(\theta)) = & - \int_0^T \left( \sum_{i=1}^d \mu_i(t; \theta) \right) dt - \sum_{n=1}^N \int_{t_n}^T \left( \sum_{i=1}^d \gamma_{i, w_n}(t - t_n; \theta) \right) dt \\ & + \sum_{n=1}^N \log \left( \mu_{w_n}(t_n; \theta) + \sum_{t_m < t_n} \gamma_{w_n, w_m}(t_n - t_m; \theta) \right) \end{aligned} \quad (7.20)$$

参数  $\theta$  可通过最大化对数似然估计。然而，数值优化是有问题的作为对数似然函数通常是相当平坦的，并可能有很多局部最大值。

#### 7.4.2.3 期望最大法

对于线性霍克斯过程中，估计也可以通过 EM 所做。EM 是 MLE 的变体，其中所述数据的一部分缺失。在分支结构表示，缺少的数据是，产生的后代的父辈。让  $z_n$  表示  $t_n$  的父辈的指数， $w_z$  代表了  $t_n$  的父的类型。如果  $z_n = m$ ， $w_z = j$ ，这代表了  $t_n$  产生了类型  $j$  的点  $t_m$ 。当  $z_n$  是 0 时， $t_n$  是一个移民。我们定义  $w_0 = 0$ ， $\gamma_{i,0}(t) = \mu_i(t)$ ， $t_0 = 0$  来简化表达式。假设  $\{t_n, w_n, z_n\}$  是已知的，因为每一代是一个独立的泊松过程中，完整的对数似然是。

$$\begin{aligned} \log(L(\theta)) = & \sum_{n=0}^N \sum_{i=1}^d \left\{ - \int_{t_n}^T \gamma_{i,w_n}(t - t_n; \theta) dt \right. \\ & \left. + \sum_{t_m > t_n} \log(\gamma_{i,w_n}(t_m - t_n; \theta)) \mathbb{1}(z_m = n) \mathbb{1}(w_m = i) \right\} \end{aligned} \quad (7.21)$$

$$\begin{aligned} Q(\theta|\theta^{(k)}) &= \mathbb{E}^{\theta^{(k)}}(\log(L(\theta))|\{(t_k, w_k)\}) \\ &= \sum_{n=0}^N \sum_{i=1}^d \left\{ - \int_{t_n}^T \gamma_{i,w_n}(t - t_n; \theta) dt + \sum_{t_m > t_n} \log(\gamma_{i,w_n}(t_m - t_n; \theta)) \right. \\ &\quad \left. \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i) \right\} \end{aligned} \quad (7.22)$$

$$\mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i) = \frac{\gamma_{i,w_n}(t_m - t_n; \theta^{(k)}) \mathbb{1}(w_m = i)}{\sum_{l=0}^{m-1} \gamma_{i,w_l}(t_m - t_l; \theta^{(k)})} \quad (7.23)$$

The EM algorithm can be implemented as follows:

1.  $k = 0$ , 选择一个初始猜测  $\theta^{(0)}$ .
2. E-step: 计算  $Q(\theta|\theta^{(k)}) = \mathbb{E}^{\theta^{(k)}}(\log(L(\theta))|\{(t_k, w_k)\})$ .
3. M-step: 计算  $\theta^{(k+1)} = \operatorname{argmax}_{\theta} Q(\theta|\theta^{(k)})$ .
4.  $K=k+1$ , 重复 E-step 和 M-step, 直到  $\theta^{(k)}$  收敛。

一般情况下, 需要以 M-步优化进行数值求解, 但是当衰减内核的指数形式  $\alpha_{ij}\beta_{ij} \exp(-\beta_{ij}t)$ , 奥尔森和卡利[39]提出一个封闭的形式近似迭代。

$$\mu_i^{(k+1)} = \frac{\sum_{m=1}^N \mathbb{P}^{\theta^{(k)}}(z_m = 0 | \{(t_k, w_k)\}) \mathbb{1}(w_m = i)}{\sum_{n=1}^N \sum_{m=n+1}^N \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i, w_n = j)} \quad (7.24)$$

$$\alpha_{ij}^{(k+1)} = \frac{\sum_{n=1}^N \sum_{m=n+1}^N \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i, w_n = j)}{\sum_{n=1}^N \sum_{m=n+1}^N \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i, w_n = j)} \quad (7.25)$$

$$\beta_{ij}^{(k+1)} = \frac{\sum_{n=1}^N \sum_{m=n+1}^N \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i, w_n = j)}{\sum_{n=1}^N \sum_{m=n+1}^N (t_m - t_n) \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i, w_n = j)} \quad (7.26)$$

$$\begin{aligned}
& \mathbb{P}^{\theta^{(k)}}(z_m = n | \{(t_k, w_k)\}) \mathbb{1}(w_m = i, w_n = j) \\
&= \frac{\alpha_{ij}^{(k)} \beta_{ij}^{(k)} \exp(-\beta_{ij}^{(k)}(t_m - t_n)) \mathbb{1}(w_m = i, w_n = j)}{\mu_i^{(k)} + \sum_{l=1}^{m-1} \alpha_{i,w_l}^{(k)} \beta_{i,w_l}^{(k)} \exp(-\beta_{i,w_l}^{(k)}(t_m - t_l))} \quad (7.27)
\end{aligned}$$

$$\begin{aligned}
& \mathbb{P}^{\theta^{(k)}}(z_m = 0 | \{(t_k, w_k)\}) \mathbb{1}(w_m = i) \\
&= \frac{\mu_i^{(k)} \mathbb{1}(w_m = i)}{\mu_i^{(k)} + \sum_{l=1}^{m-1} \alpha_{i,w_l}^{(k)} \beta_{i,w_l}^{(k)} \exp(-\beta_{i,w_l}^{(k)}(t_m - t_l))} \quad (7.28)
\end{aligned}$$

此外， $\exp(-\beta_{i,w_l}^{(k)}(t_m - t_l))$  衰减到一个较小的值后的总和，

$\sum_{l=1}^{m-1} \alpha_{i,w_l}^{(k)} \beta_{i,w_l}^{(k)} \exp(-\beta_{i,w_l}^{(k)}(t_m - t_l))$  可被截断。EM 的速度被报道为快 10-100 倍 MLE 和，更重要的是，最大似然估计并不能在 500 次迭代内的所有的测试案例收敛，而 EM 可以。

## 7.5 霍克斯过程的应用

基本理论和霍克斯过程统计推断的基础之后，我们现在用其去模拟高频数据模型。首先，读者知道股票交易频率的概念如何不同即可。根据 2014 年 5 月 2 日的交易和报价数据库上午 9:30 至下午 4:00，有 1100 万的报价变化（上限+取消+市场订单）和 0.3 万宗交易（市场订单）的 SPDR S&P 500 指数 ETF（SPY）。换言之，平均有 460 报价的变化和每秒 13 交易。如果我们把一个快照每隔 5 分钟，我们将使用只有 0.03% 的交易数据和报价数据的 0.0007%。相比之下，探路者 Bancorp 的仅具有 306 报价的变化和图 11 的交易的同一天，这意味着有平均交易因而 5 分钟快照之间有 35 分钟的滞后将只给予一系列的重叠的信息。无论采样频率，我们很可能，如果我们使用固定的时间间隔，流动性好和流动性差的股票投资组合分析异步数据得到一些误导性的结果。

多元点进程的结构表明，每变动内容可以具有一个完全不同的到达强度  $\mathbf{l}(t)$  的。尽管如此，多元霍克斯过程仍然可以依赖结构容易经由其通过适当考虑的最高频率的所有异步数据无需任何重采样估计  $\gamma_{ij}(t)$  的建模。

为了到达和价格的变化是无可争议的两个高频交易中最重要元素。霍克斯使用过程中，我们可以估算的基础上所有的历史高频异步数据的条件分布，使我们能够给未来事件发生的更为准确的实时预测。在接下来的小节中，我们将重点介绍一些这需要霍克斯过程的优势，高频数据模型中的文章

### 7.5.1 订单到达建模

**Bowsher** 是第一个使用霍克斯过程来为订单到达建模的人。他使用非线性霍克斯进程以实现抑制效果，而他认为两率函数  $\Phi_i(x, t) = \mu_i(t) + \exp(x)$  和  $\Phi_i(x, t) = \max(\mu_i(t) + x, \varepsilon_i)$ ,  $\varepsilon_i > 0$ ，两式都保证随机强度会在任何时候都严格为正。对于确定基准比  $\mu_i(t)$ ，他利用与节 9:30, 10:00, 11: 00, ..., 16 分段线性函数，而衰减内核是没有痕迹的指数函数。其中额外的一项表示从前一交易日的溢出效应。

**Bowsher** 用 MLE 来估计贸易和通用汽车的报价二元点过程中的参数，2000 年 7 月 5 日和 8 月 29 日至 2000 年在纽约证券交易所交易模型发现根据使用随机时间的变化所得结果拟合度高。

**Large** 采用霍克斯过程只激进的订单，而不是所有的交易和报价的模型。这是市场的订单消耗队列和限价的买卖价差内跌倒的到达模型，以研究弹性的 **LOB**。如果经过大交易将恢复其通用形状的 **LOB** 被称为弹性。我们的想法是，当一个大的贸易会导致买卖价差扩大，积极的限价订单的到达强度的弹性 **LOB** 将激增，这样的差距将很快被填补。为了也就是说，从积极的市场订单，以积极的限价订单交激励效应  $\gamma_{ij}(t)$  应该是一个弹性相当大的 **LOB**。

除了市场订单和限价订单，**Large** 还考察了限价订单的取消以及限价订单的最佳报价外跌落。因此，他建立了一个 10-线性变动内容霍克斯标记与过程指数衰减和标记冲击内核，以适应巴克莱银行 (**BARC**) 的高频数据，1 月 2 日和 31 日之间的交易 **LSE**，2002 年的结果表明，性招标的扩大差价确实使进取限价订单的强度升高，造成的差距非常迅速，从而填补使得 **LOB** 变得有弹性。

### 7.5.2 价格跳跃建模

#### 7.5.2.1 单项资产

传统上，价格跳跃的事件由泊松过程，带有引言部分中提到的缺点。但霍克斯过程可以应用于模型价格跳跃，这往往划定聚类，自和交叉激励的行为。

**Bacry** 等使用霍克斯过程的价格跳跃建模，导致能够再现的微观结构噪声，艾普效果，以及跳聚类模型，同时保持布朗扩散的粗尺度限制。在他们的模型，交易价格  $X(t)$  具有动态

$$X(t) = N_1(t) - N_2(t) \quad (7.29)$$

其中， $N(t) = (N_1(t), N_2(t))$  是指数衰减内核二元线性霍克斯过程。 $N_1(t), N_2(t)$  表示向上和向下的跳跃的总数分别。作者提出额外的假设，该霍克斯  $N$  个进程只有交叉激发和系数是对称的，以简化计算。

$$\begin{aligned} \lambda_1(t) &= \mu + \int_{(0,t)} \gamma(t-s) dN_2(s), \quad \lambda_2(t) = \mu \\ &+ \int_{(0,t)} \gamma(t-s) dN_1(s), \quad \gamma(t) = \alpha e^{-\beta t} \end{aligned} \quad (7.30)$$

根据该模型，当  $X$  跳跃上升（下降）， $\lambda_2$ （相应地， $\lambda_1$ ）的增加，导致跳向下（相应向上），以增加的概率。这样的交联生成交易价格是最佳的出价之间弹跳和最佳问显微噪声的影响。

因为买卖反弹，这是众所周知的实现的方差增加的采样频率增加时间。

$$V(\tau) = \mathbb{E} \left( \frac{1}{T} \sum_{n=0}^{T/\tau} (X((n+1)\tau) - X(n\tau))^2 \right) \quad (7.31)$$

$$= \frac{2\mu}{1 - \alpha/\beta} \left( \frac{1}{(1 + \alpha/\beta)^2} + \left( 1 - \frac{1}{(1 + \alpha/\beta)^2} \right) \frac{1 - e^{-(\alpha+\beta)\tau}}{(\alpha + \beta)\tau} \right) \quad (7.32)$$

这样的效果可以通过计算跳转模型（7.29）和结果的预期实现方差（7.32）。很容易地证明  $\mu = 0.16$ ,  $\alpha = 0.024$ ,  $\beta = 0.11$ , 如图 7.1 所示。作者运用该模型欧洲外滩期货, 找到这个高度简化的模型下的观察和理论认识的差异之间一个非常不错的选择。

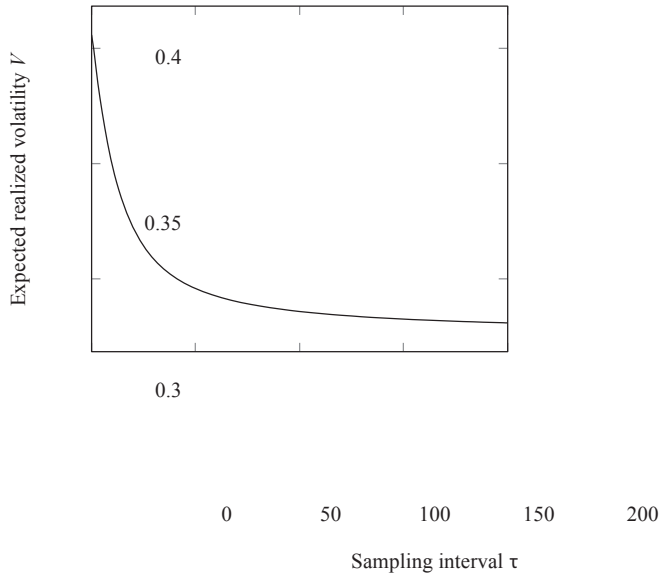


FIGURE 7.1 Volatility signature plot of Hawkes jump model.

表明宏观霍克斯跳跃模式可以追溯到布朗运动的经典模型, 由于功能性中心极限定理线性霍克斯过程（7.7）。假设  $\int_0^\infty \gamma(t) dt < 1$

$$\frac{X(n\bullet)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\text{weak}} \sigma W(\bullet), \quad \sigma^2 = \frac{2\mu}{(1 - \int_0^\infty \gamma(t) dt)(1 + \int_0^\infty \gamma(t) dt)^2} \quad (7.33)$$

有趣的是看到了宏观方差  $\sigma^2$  是如何相关的微观基率  $\mu$  和交叉激励  $\gamma(t)$  的。当  $\int_0^\infty \gamma(t) dt$  接近 1, 方差变为  $\infty$ 。

Jaisson and Rosenbaum 延伸 Bacry 等人的模型至接近不稳定霍克斯处理的情况下, 其中  $\int_0^\infty \gamma(t) dt \simeq 1$ , 通过构造的核函数, 其积分收敛到 1 在第  $n-1$  的速度的顺序。他们表明, 适当的比例定价过程收敛与赫斯顿随机波动率布朗扩散。完整的结果如下所述。



$$X^{(n)}(t) = N_1^{(n)}(t) - N_2^{(n)}(t) \quad (7.34)$$

$$\lambda_1^{(n)}(t) = \mu + \int_0^t \gamma_1^{(n)}(t-s) dN_1^{(n)}(s) + \int_0^t \gamma_2^{(n)}(t-s) dN_2^{(n)}(s) \quad (7.35)$$

$$\lambda_2^{(n)}(t) = \mu + \int_0^t \gamma_2^{(n)}(t-s) dN_1^{(n)}(s) + \int_0^t \gamma_1^{(n)}(t-s) dN_2^{(n)}(s) \quad (7.36)$$

$$\left( \int_0^\infty \gamma_1^{(n)}(t) dt + \int_0^\infty \gamma_2^{(n)}(t) dt \right) < 1, \quad \gamma_i^{(n)}(t) = \alpha^{(n)} \gamma_i(t) \quad (7.37)$$

$$\gamma_i : \mathbb{R}_+ \longrightarrow \mathbb{R}_+, \quad \int_0^\infty (\gamma_1(t) + \gamma_2(t)) dt = 1, \quad \int_0^\infty t(\gamma_1(t) + \gamma_2(t)) dt = m < \infty \quad (7.38)$$

$$\int_0^\infty |\gamma_i'(t)| dt < \infty, \quad \sup_{t \in [0, \infty)} |\gamma_i'(t)| < \infty \quad (7.39)$$

$$\alpha^{(n)} \in [0, 1), \quad \lim_{n \rightarrow \infty} \alpha^{(n)} = 1, \quad \lim_{n \rightarrow \infty} n(1 - \alpha^{(n)}) = c > 0 \quad (7.40)$$

$$\begin{aligned} \psi^{(n)}(t) &= \sum_{k=1}^{\infty} \left( \gamma_1^{(n)} + \gamma_2^{(n)} \right)^{\otimes k}(t), \quad \rho^{(n)}(t) = \frac{n\psi^{(n)}(nt)}{\int_0^\infty \psi^{(n)}(t) dt}, \\ |\rho^{(n)}(t)| &\leq M \quad \forall n \quad \forall t \end{aligned} \quad (7.41)$$

$$\frac{X^{(n)}(n\bullet)}{n} \xrightarrow[n \rightarrow \infty]{\text{weak}} Y(\bullet) \quad (7.42)$$

$$dY_t = \frac{\sqrt{V_t}}{1 - \int_0^\infty |\gamma_1(t) - \gamma_2(t)| dt} dW_t^1, \quad Y_0 = 0 \quad (7.43)$$

$$dV_t = \frac{c}{m} \left( \frac{2\mu}{c} - V_t \right) dt + \frac{\sqrt{V_t}}{m} dW_t^2, \quad V_0 = 0 \quad (7.44)$$

条件 (7.34-7.36) 是一个二元霍克斯模型 (自我和交叉激励), 但现在我们有用来缩

放时每个  $n$  不同的  $\gamma_i^{(n)}(t)$ 。其余的都是类似的单变量近不稳定霍克斯过程 (7.17)

的规律性条件和最重要的一项是 (7.40), 其中指出,  $\alpha(N)$  趋近于 1 为  $n-1$  的速度。然而, 有趣的结果是代替会聚到一个集成的 CIR, 由两个霍克斯过程之间的差异形成的价格动态收敛到一个随机波动模型。

#### 7.5.2.2 两项资产



为了给 Epps effect 建模, Bacry 设想了一个两项资产的情况, 其中价格  $(X_1(t), X_2(t))$  由下给出

$$X_1(t) = N_1(t) - N_2(t), \quad X_2(t) = N_3(t) - N_4(t) \quad (7.45)$$

$$\lambda_i(t) = \mu_i + \sum_{j=1}^4 \int_{(0,t)} \alpha_{ij} \exp(-\beta(t-s)) dN_j(s), \quad i = 1, \dots, 4 \quad (7.46)$$

$(N_1(t), \dots, N_4(t))$  是一个带有指数形式的衰退核的 4 变量霍克斯过程其中  $\beta_{ij} = \beta$ , 激发作用的耦合被约束为以下形式

$$\alpha = \begin{pmatrix} 0 & \alpha_{12} & \alpha_{13} & 0 \\ \alpha_{12} & 0 & 0 & \alpha_{13} \\ \alpha_{31} & 0 & 0 & \alpha_{34} \\ 0 & \alpha_{31} & \alpha_{34} & 0 \end{pmatrix} \quad (7.47)$$

在这种情况下, 有一个封闭的形式表示为实现相关性, 当采样间隔变为 0, 该相关性消失。如果假设  $\mu_1 = \mu_2, \mu_3 = \mu_4, \alpha_{12} = \alpha_{34} = 0, \left(\int_0^\infty \gamma_{13}(t) dt\right)$

$\left(\int_0^\infty \gamma_{31}(t) dt\right) < 1$ , 那么宏观二元资产价格收敛于相关布朗扩散

$$\frac{1}{\sqrt{n}} \begin{pmatrix} X_1(n\bullet) \\ X_2(n\bullet) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{\text{weak}} \frac{\sqrt{2} \begin{pmatrix} \sqrt{v_1} W_1(\bullet) + \sqrt{v_2} \int_0^\infty \alpha_{13}(t) dt W_2(\bullet) \\ \sqrt{v_1} \int_0^\infty \alpha_{31}(t) dt W_1(\bullet) + \sqrt{v_2} W_2(\bullet) \end{pmatrix}}{(1 - (\int_0^\infty \gamma_{13}(t) dt)(\int_0^\infty \gamma_{31}(t) dt))^{3/2}} \quad (7.48)$$

$$v_1 = \mu_1 + \left(\int_0^\infty \gamma_{13}(t) dt\right) \mu_3, \quad v_2 = \mu_3 + \left(\int_0^\infty \gamma_{31}(t) dt\right) \mu_1$$

( $W_1, W_2$ ) is standard 2-dimensional Brownian motion (7.49)

这种融合的结果给了我们一个明确的公式来估算从异步高频数据的宏观相关性。

作为最后的一句话, 这跳跃的表征下, 观察到的交易价格是不是有些隐藏的连续公允价值的过程, 加上一些微观结构噪音。它是在一个固定的价格格买家 ( $N_1, N_3$ ) 和卖方 ( $N_2, N_4$ ) 之间的交易的相互作用的结果。有作为高频波动或关联没有这样事情, 因为价格不扩散, 但纯粹的跳跃过程中的高频规模。当我们看规模扩散近似时波动性和相关性是唯一有意义的, 但这些低频表现参数可以直接从高频跳跃模式的特点进行计算。

### 7.5.3 跳跃扩散建模

Duffie 提出仿射(affine)跳扩散  $X(t)$ , 其具有以下结构。

$$dX(t) = (k_0(t) + k_1(t)X(t))dt + (h_0(t) + h_1(t)X(t))dW(t) + \zeta dN(t) \quad (7.50)$$

$$\lambda(t) = a_0(t) + a_1(t)X(t) \quad (7.51)$$

$N(t)$  的跳跃强度  $\lambda(t)$  是  $X(t)$  的一个仿射函数, 这个函数依赖于伯努利运动  $W(t)$  和跳跃过程  $N(t)$ , 带有  $\zeta$  从固定的分布绘制。当  $k_0 = \beta\theta, k_1 = -\beta, h_0 = h_1 = 0, a_0 = 0, a_1 = 1, \zeta = \alpha$ , 我们可以得出  $\lambda(t) = X(t)$  and  $d\lambda(t) = \beta(\theta - \lambda(t))dt + \alpha dN(t)$ 。所以在这种情况下,  $N(t)$  是一个带有指数形式的衰退核的霍克斯过程。

在  $\zeta$  是一个常数的条件下, Zhu 推导出了一些收敛结果, 所述扩散部分是一个 CIR 过程。

$$dX(t) = \beta(\mu - X(t))dt + \sigma\sqrt{X(t)}dW(t) + \alpha dN(t) \quad (7.52)$$

$$\lambda(t) = a_0 + a_1X(t) \quad (7.53)$$

Aït-Sahalia 等人用霍克斯跳跃扩散给金融危机的传染性建模, 其中价格动态  $X_i(t)$  由下给出

$$dX_i(t) = \mu_i dt + \sqrt{V_i(t)}dW_i^X(t) + Z_i(t)dN_i(t) \quad (7.54)$$

$$dV_i(t) = \kappa_i(\theta_i - V_i(t))dt + \eta_i\sqrt{V_i(t)}dW_i^V(t) \quad (7.55)$$

扩散部分是 Heston 随机波动率模型和跳转的部分是一个多元霍克斯过程多资产之间跳跃的聚集和传播。 $Z_i(t)$  的对应于跳跃大小和方向。

### 7.5.4 测量内生性

介于霍克斯分支结构代表了时间到达, Filimonov 描绘移民为外生的消息, 而后代是内生的事件。在股票市场价格变动的背景下, 移民是价格发现, 因知情从而对外部信息反应制定订单的交易者, 而后代是由噪声交易者, 那些追随而产生的不稳定涟漪, 势头交易, 寄生虫交易等。

根据与指数衰减内核和恒定基准利率元线性模型霍克斯, 每个个体直接后裔的预期数量由下式给出

$$n = \int_0^\infty \gamma(s)ds = \int_0^\infty \alpha e^{-\beta s} ds = \alpha/\beta \quad (7.56)$$

对于给定的移民, 后代在所有后代的预期数量是  $n + N_2 + N_3 + \dots = N/(1-n)$  的如果  $n < 1$ , 所以后代与总人口的比例 (非移民) 是

$$\frac{\text{Descendants}}{\text{Descendants} + \text{immigrant}} = \frac{n/(1-n)}{n/(1-n) + 1} = n \quad (7.57)$$

因此, 分支系数  $n$  特征化内生性反馈活动的数量, 而基准比率  $\mu$  测量外源性信息到达率。

使用 E-mini S&P 特征作为代表 Filimonov 和 Sornette 发现美国市场中的内生性

水平  $n$  已经从 1998 年的 0.3 升到 2007 年的 0.7. 此外, 在 5 月 6 日闪电崩盘,  $n$  到达了 0.95 的峰值。