微观视角下的波动率：产生、度量与影响

几乎处处学不会

2024.3.24

1 概要

对于各种尺度下波动率的计算大家并不陌生，从高频尺度下的 RV 到低频尺度下的 CTC，Parkinson，GK，Yang-Zang 等，在统计上我们有各种各样的公式可以用来计算波动率，他们在不同的场景下各有优劣。但是，这些波动率的计算方法更多是站在统计学的视角来进行估计的，而无法直接体现波动率的产生机制。本文将从较为微观的视角出发，介绍 3 个模型，以探讨波动率的产生机制，度量偏误以及对市场的影响。

2 波动率的产生

在本章，我们将通过建模揭示波动率和交易量及信息密度的关系。我们知道在微观尺度下，市场中存在 taker 和 maker 两类交易者，taker 根据报价交易，maker 提供报价。为了方便建模，我们假设：

1. 信息由 taker 的交易行为传播，maker 仅根据成交被动调整价格
2. 买卖无价差
3. 买卖单的发生在单位时间内服从泊松流

在假设条件下，我们定义价格对交易量的响应函数为 $F(n_i) = sgn(n_i) * \alpha * \sqrt{|n_i|}$，而在 Δt 时刻内价格的变化为：

$$\Delta \ln(s) = \sum_{i=1}^{N(\Delta t)} F(n_i) \tag{1}$$

其中 $N(\Delta t)$ 表示在 Δt 时刻内订单到达的次数，n_i 表示第 i 次订单的数量。接下来我们利用条件期望和方差来计算波动率：

$$\begin{aligned}
\sigma_t^2 = Var(\Delta \ln(s)) &= E[Var(\Delta \ln(s)|N(\Delta t))] + Var(E[\Delta \ln(s)|N(\Delta t)]) \\
&= E[N(\Delta t) * Var(F(n_i))] + Var(N(\Delta t) * E[F(n_i)]) \\
&= E[N(\Delta t)] * Var(F(n_i)) + Var(N(\Delta t)) * (E[F(n_i)])^2 \\
&= \lambda \Delta t * (Var(F(n_i)) + (E[F(n_i)])^2) = \lambda \Delta t * E[(F(n_i))^2] = \alpha^2(\lambda \Delta t E(|n|)) = \alpha^2 \mu_t
\end{aligned} \tag{2}$$

其中 μ_t 表示交易量，α 会是一个反应交易弹性的一个参数，他会是一个和信息密度、流通形式等因素有关的参数，对于商品期货，它还会是一个和库存水平有关的参数。

通过上式我们揭示了，波动率会和交易量保持一定的正相关关系。虽然这是一个较为简单的模型，但是它在具备一定的合理性下为我们提供了一个从微观尺度下理解波动率的思路。当然，在现实中，如果假设 1 不被满足，意味着信息不完全借由交易量传递，也就是说，maker 也会根据接收到的信息主动调整价格，届时我们假定的响应函数将不成立，价格直接由市场共识形成，本模型将不再适用，一个更直观的情形就是当一个重大事件发生时，maker 甚至不需要 taker 的交易行为，就会判断出现在的价格不合理，从而主动调整价格，这种情况下交易量变化不大，意味着 α 会很小，那么此时模型将失效。

另一个在实务中的重要问题是参数 α 的确定，我们可以由历史数据做回归来得到一个 α 的估计，但是中间要对 X 和 Y 做一些非线性变换（我们要注意的有非线性变换的形式，以及 X 和 Y 具体应该选取什么指标，采样的频率等），并且回归的显著性也并非总是很高，在实务中如何处理不显著的情形，也是一个需要考虑的地方。

3 波动率的度量

在最开始的概要中，我们提到了各种波动率的度量方法，我们也指出，波动率的估计常常是有偏误的，尤其是在高频尺度下，那么，这种偏误是如何产生的呢？接下来，我们将借用 ROLL 模型来讨论波动率的估计偏误的产生原因。ROLL 模型主要基于以下假设：

1. 信息由 maker 的行为反应，maker 会根据信息主动调整价格
2. taker 行为随机，不含信息量

在这个假设前提下，我们认为标的价格 p_t 满足：

$$p_t = m_t + q_t * c \tag{3}$$

其中 m_t 表示有效价格（中间价），$c = \frac{tick}{2}$，而 q_t 满足：

$$q_t = \{ \begin{matrix} 1, \text{buy} \\ -1, \text{sell} \end{matrix} \tag{4}$$

由于 taker 行为随机且不含信息量，我们可以认为 q_t 满足：$P(q_t = 1) = P(q_t = -1) = \frac{1}{2}$，并且 q_i、q_j 和 m_t 是独立的，同时 $E[\Delta p_t] = 0$，那么我们可以得到：

$$\begin{aligned} Var[\Delta p_t] &= E[(\Delta p_t)^2] \\ &= E[((m_t - m_{t-1}) + (q_t - q_{t-1}) * c)^2] = E[(u_t + (q_t - q_{t-1}) * c^2)] \\ &= E[u_t^2 + 2u_t(q_t - q_{t-1}) * c + q_t^2 * c^2 + q_{t-1}^2 * c^2 - 2q_t * q_{t-1} * c^2] \\ &= E[u_t^2] + 0 + c^2 E[q_t^2] + c^2 E[q_{t-1}^2] - 0 \\ &= \sigma_u^2 + 2c^2 \end{aligned} \tag{5}$$

上式中我们用到了 q_t 和 m_t 及 q_{t-1} 的独立性，以及 $E[q_t^2] = E[q_{t-1}^2] = 1$。由此我们可以看出，波动率包含了买卖价差 c 的信息，因此当买卖价差较大的时候，c 较大，波动率也会较大，但我们往往在计算波动率的时候只使用 K 线数据，意味着我们只考虑了有效价格 m_t 对波动率的贡献 σ_u^2 ，而没有考虑买卖价差 c，这就是在高频尺度下波动率估计的偏误的一部分原因。

除去上述原因，价格序列的负相关性也是造成波动率估计有偏的原因之一，为此我们计算序列和滞后 1 期的序列的协方差，由于和式 5 的计算过程雷同，我们在此省略过程，得到结果为：

$$Cov(\Delta p_t, \Delta p_{t-1}) = Cov(u_t + (q_t - q_{t-1}) * c, u_{t-1} + (q_{t-1} - q_{t-2}) * c) = -c^2 \tag{6}$$

由此我们可以看出，价格序列的负相关性使得独立同分布的假设不成立，从而使得波动率估计有偏。

4 波动率的影响

我们在前面讨论了波动率的产生和度量，接下来我们来讨论波动率对市场的影响。这一部分我们将采用 Glosten and Milgrom(1985) 的模型来进行讨论。我在这里仅作简单推导，具体的推导可以参考材料 1。Glosten and Milgrom 的模型中

首先假设价格在未来有两种可能，V_H 和 V_L 对应涨和跌，当前 LOB 中有两个价格，卖方价格 a 和买方价格 b；其次，市场中有两类交易者，第一类是知情交易者，他们知道未来会是 V_H 还是 V_L，所以当未来会是 V_H 并且 $a < V_H$ 的时候他们会买入（这个概率会是 p），反之卖出；第二类是不知情的流动性交易者，他们不知道未来会是 V_H 还是 V_L，所以他们的买卖行为会是等概率的。市场中知情者的比例是 α，这些知情者使得买卖力量不均衡，对应的 maker 需要根据知情者比例来调整价格，Glosten and Milgrom 的模型指出，当做市商围绕着中间价做市时，以 ask 为例子，他的报价 $\mu + Delta_a$ 的期望收益是：

$$\frac{(1-\alpha)/2}{\alpha p + (1-\alpha)/2}\Delta_a + \frac{p\alpha}{\alpha p + (1-\alpha)/2}\left(\Delta_a - (V_H - \mu)\right) \tag{7}$$

合理的报价应该使得上述期望收益为 0，进一步，对 bid 同理，最后，买卖关于中间价的理论价差应该是

$$\Delta_a = \frac{\alpha p}{\alpha p + (1-\alpha)/2}\left(V_H - \mu\right) \tag{8}$$

及

$$\Delta_b = \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{1/2}{1-p}}\left(\mu - V_L\right). \tag{9}$$

这表明买卖价差除去和知情者比例 α 相关外，还和未来价格的可能取值 V_H 和 V_L 有关，后者会和波动率息息相关，当波动率较大的时候，未来价格的分布方差较大，意味着 V_H 和 V_L 分别和中间价的期望 $\mu = E[v|\mathcal{F}]$ 差距会更大，进一步使得买卖价差更大，这便是波动率对微观尺度下的市场的影响之一。

5 结论

可以看出基于上述的讨论，信息通过 taker，由交易量影响波动率，而波动率会反过来影响 maker 调整买卖价差，而买卖价差又会进一步影响 taker 的决策，因此波动率在其中起到了一个桥梁的作用，让信息在 maker 和 taker 之间充分利用，从而使得市场价格更加合理。然而，正如我们前文所讨论的，一些特殊的结构下可能会使得上述模型失效。因此在实务中我们还需进行更精细的建模和分析。

6 参考材料

1. Algorithmic and High-Frequency Trading (Álvaro Cartea, Sebastian Jaimungal etc.)

2. A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market (Roll, 1984)