Separation of Diffusive and Jump Returns

1 The Basic Setup

Our main model for now is
\[ dX_t = \sqrt{c_t} \, dW_t + J_t \]  
(1)
where \( c_t \) is the local variance and \( J_t \) is a centered jump process of finite activity.

The raw price data are equally spaced \( P_{k\Delta_n}, k = 0, 1, 2, \ldots, nT \). The convention is always that the unit of time measurement is one day, \( n \) counts the number of within day observations and \( T \) counts the number of days. We form daily within-day returns
\[ \Delta^n_{(i-1)n+1} X, \quad i = 1, 2, \ldots, n \quad X = 100 \times \log(P) \]  
(2)
and we will write \( X = X^c + X^d \) to separate diffusive and jump parts. In (2) we do not keep differences in prices across days because of overnight returns, and we thus we keep \( n \) within-day returns for \( T \) days. The daily integrated and jump variations are
\[ IV_t = \int_{t-1}^t c_s \, ds, \quad JV = \sum_{t-1<s \leq t} \Delta^2 J_s. \]  
(3)

For less cumbersome notation, let integer \( t = 1, 2, \ldots, T \) index days and define the returns by
\[ r_{t,i} = \Delta^n_{(i-1)n+1} X \]  
(4)
The realized and bipower variations are
\[ RV_t = \sum_{i=1}^n r_{t,i}^2 \]  
(5)
\[ BV_t = \pi \frac{n}{2n-1} \sum_{i=2}^n |r_{t,i}| |r_{t,i-1}| \]  
(6)
The term \( \frac{n}{n-1} \) is a small degree of freedom adjustment which can be import when \( n \) is small.

2 Separating Diffusive and Jump Moves

We want to classify the individual moves as diffusive or jump. If we observed (we do not) the variance path \( c_s \) for \( s \in [t_i - \Delta_n, t_i] \) this would be relatively easy. Let \( \sigma_{ti} = \sqrt{\int_{t_i-\Delta_n}^{t_i} c_s \, ds} \) denote the local standard deviation over intervals \([t_i - \Delta_n, t_i]\). Under the model and notation
\[ r_{t,i} = r_{t,i}^c + r_{t,i}^d \]  
(7)
The fluctuations in the continuous part are
\[ \Delta_n^\circ X \approx \sigma_{ti} Z_i \]  
(8)
where \( \sigma_{ti} \) is the local volatility and \( Z_i \sim N(0, 1) \). A rule like
\[
|r_{i,t}| \leq \alpha_n \sigma_{ti} \Rightarrow \text{diffusive move} \quad |r_{i,t}| > \alpha_n \sigma_{ti} \Rightarrow \text{jump move},
\]  
(9)
where \( \alpha_n \to 0 \) as \( \Delta_n \to 0 \) appropriately, should correctly disentangle the diffusive and jump moves. This sort of thresholding is due to Mancini (2001, 2009) and refined in Jacod and Protter (2012). The fact that something like the above actually correctly classifies moves is a sort of folk theorem that seems to be implicit in Jacod and Protter (2012). A formal proof is given in Li et al. (2015).

The classification is done as
\[
|r_{ti}| \leq \text{cut}_{ti} \rightarrow \text{diffusive} \quad |r_{ti}| > \text{cut}_{ti} \rightarrow \text{jump}
\]  
(10)
To set the cut value \( \text{cut}_{ti} \) we let \( BV_t \) denote the overall level of volatility on day \( t \). Then to account for the daily (diurnal) pattern of volatility, we estimate the mean (across days) of the bi-power factors:
\[
b_i = \frac{1}{T} \sum_{t=1}^{T} |r_{ti}| |r_{t,i-1}| \quad i = 2, 3, \ldots, n
\]  
(11)
and set \( b_1 = b_2 \). The time-of-day (tod) factor is
\[
\tau_i = \frac{b_i}{\frac{1}{n} \sum_{j=1}^{n} b_j}.
\]  
(12)
Above, we re-scale \( b_i \) so that the mean over the day is unity. The cut factors are then
\[
\text{cut}_{ti} = \alpha \sqrt{\tau_i BV_t \Delta_n^{0.49}}
\]  
(13)
Our current practice is to separate as follows:
\[
r_{t,i}^c = r_{t,i} 1_{|r_{t,i}| \leq \alpha \Delta_n^{0.49} \sqrt{\tau_i BV_t}}
\]  
(14)
\[
r_{t,i}^d = r_{t,i} 1_{|r_{t,i}| > \alpha \Delta_n^{0.49} \sqrt{\tau_i BV_t}}
\]  
(15)
where \( \alpha \in [3.5, 4.5] \) is a constant equal to a relatively large number of local standard deviations. Keep in mind that the diurnal pattern \( \tau_i \) redistribute the total variation \( IV_t \) over the day. We use \( BV_t \), the same day’s bi-power variation, as a proxy for \( IV_t \). Since \( \Delta_n^{0.49} \) equals \( \Delta_n^{0.50} \times \Delta_n^{-0.01} \) it is seen that that the band decreases more slowly than the rate \( \Delta_n^{0.50} \).

Another view is this: The local variance is roughly constant and proportional to the total daily diffusive variance \( IV_t \) (proxied by \( BV_t \)). Thus the fluctuations in the returns are like
\[
\text{local return} \approx \Delta_n^{0.50} \sqrt{\tau_i IV_t Z_i}
\]  
(16)
That is, it as if they local return is $N(0, \Delta_n \tau_i IV_t)$. We need the cutoff to decrease slower $\Delta_n^{0.50} \sqrt{\tau_i IV_t}$ in order to let all of the diffusive moves get through but exclude the discontinuous jump moves. Thus we use a cutoff $\Delta_n^{0.49} \sqrt{\tau_i IV_t}$. Note that

$$\frac{\Delta_n^{0.49} \sqrt{\tau_i BV_t}}{\Delta_n^{0.50} \sqrt{\tau_i IV_t}} \approx \Delta_n^{-0.01} \to \infty$$  \hspace{1cm} (17)

In summary, we use BV to approximate today’s total diffusive variance and cut on the basis

$$\Delta_n^{0.49} \sqrt{\tau_i BV_t}$$  \hspace{1cm} (18)

instead of using $\Delta_n^{0.50}$. Actually, we could use any exponent close to 0.50, and the literature often writes

$$\text{cut}_{t,i} = \Delta_n^a \sqrt{\tau_i BV_t}, \quad 0 < a < \varpi < 0.50$$  \hspace{1cm} (19)

where the lower bound $a$ is determined by some other (unimportant here) conditions. Observe that $\sum_{i=1}^n (r_{t,i}^c)^2$ is the truncated variance estimator of $IV_t$. 

3
References


