For the Poster of IMA Workshop 9, May 24 - 28, 2004:
Financial Data Analysis and Applications

Filtering with a Marked Point Process Observation:

Applications to the Econometrics

of Ultra-High-Frequency Data

Yong Zeng
Department of Mathematics and Statistics
University of Missouri at Kansas City

Outline of Study

1. **Brief Overview of modeling UHF data: three directions**
2. **Direction Three: two different views of UHF data**
   (a) First View: An Irregularly-Spaced Time Series (with review)
   (b) Second View: A Realized Sample Path of the MPP
   (c) Detailed Review of the Second View
3. **The Proposed Framework**
   (a) Filtering with a MPP observation
   (b) Construction of the Partially-Observed Model
   (c) Examples
4. **Continuous-time Likelihoods**
5. **Future Works**

Brief Overview of modeling UHF data

Direction One:

Transaction price = Intrinsic value + Microstructure noise

Tools: Vector AR impulse response analysis and state-spaced models

Recent developments: Hasbrouck (1996), (1999), (2002) and George and Hwang (2001)

Direction Two:

Aggregate UHF data into time series of five-minute intervals or others

Advantage: construct *realized volatilities*

Recent developments: Anderson, Bolleslev, Diebold and Labys (2003) and Barndorff-Nielsen and Shephard (2004)

Direction Three: MPP models

The First View: An Irregularly-Spaced Time Series (Engle (2000))
Date: $\{(\Delta t_i, y_i), i = 1, ..., N\}$
Log likelihood:

$$(\Delta t_i, y_i)|\mathcal{F}_{i-1} \sim f(\Delta t_i, y_i|\breve{\Delta} t_{i-1}, \breve{y}_{i-1}; \theta)$$

where $\breve{z}_i = \{z_i, z_{i-1}, ..., z_1\}$

$$f(\Delta t_i, y_i|\breve{\Delta} t_{i-1}, \breve{y}_{i-1}; \theta) = g(\Delta t_i|\breve{\Delta} t_{i-1}, \breve{y}_{i-1}; \theta) q(y_i|\Delta t_i, \breve{\Delta} t_{i-1}, \breve{y}_{i-1}; \theta)$$

$$L(\Delta, Y; \theta) = \sum_{i=1}^{N} \log g(\Delta t_i|\breve{\Delta} t_{i-1}, \breve{y}_{i-1}; \theta) + \sum_{i=1}^{N} \log q(y_i|\Delta t_i, \breve{\Delta} t_{i-1}, \breve{y}_{i-1}; \theta)$$

Developments under the First View

1. Autoregressive Conditional Duration (ACD) model by Engle and Russell (1998) and logarithmic ACD by Bauwens and Giot (2000),threshold ACD by Zhang, Russell and Tsay (2001), Asymmetric ACD by Bauwens and Giot (2003)
2. Bivariate Point process model by Engle and Lunde (2003), and Autoregressive Conditional Intensity model by Russell (1999)
3. Ordered Probit model by Hausman, Lo and Mackinlay (1992) Autoregressive Conditional Multinomial by Russell (1998), Activity-Direction-Size model by Rydberg and Shepard (2003), Price Change Duration model by McCulloch and Tsay (2001)
4. UHF-GARCH by Engle (2000), ACD-GARCH by Ghysels and Jasiak (1998)

The Second View: A Realized Sample Path of the MPP

Zeng (2003) and Kouritzin and Zeng (2004)

• **Modeling Prices as a Collection of Counting Processes**
or *filtering with counting process observations*

$$\vec{Y}(t) = \begin{pmatrix} N_1(\int_0^t \lambda_1(\theta(s), X(s), s)ds) \\ N_2(\int_0^t \lambda_2(\theta(s), X(s), s)ds) \\ \vdots \\ N_n(\int_0^t \lambda_n(\theta(s), X(s), s)ds) \end{pmatrix}, \tag{1}$$

where $Y_j(t) = N_j(\int_0^t \lambda_j(\theta(s), X(s), s)ds)$ is the counting process recording the cumulative number of trades that have occurred at the jth price level up to time t.

Assumption 1.1: (θ, X) *is the solution of a martingale problem for a generator* $\mathbf{A}$ *such that* $M_f(t) = f(\theta(t), X(t)) - \int_0^t \mathbf{A}f(\theta(s), X(s))ds$ *is a* $\mathcal{F}_t^{\theta,X}$*-martingale.*
Assumption 1.2: $(N_1, \dots, N_n)$ are unit Poisson processes under measure P.
Assumption 1.3: $(\theta, X), N_1, \dots, N_n$ are independent under measure P.
Assumption 1.4: $0 \le a(\theta(t), X(t), t) \le C$ for some $C > 0$ and all $t > 0$.
Assumption 1.5: Intensities: $\lambda_j(\theta, x, t) = a(\theta, x, t)p(y_j|x)$, where $a(x, \theta, t)$ is the total intensity, and $p_j = p(y_j|x)$ is the transition probability from x to y_j.

Likelihoods and Posterior

Continuous-time Joint Likelihood Function of (θ, X, Y)

$$L(t) = \frac{dP}{dQ}(t) = \prod_{k=1}^{n} \exp\left\{ \int_0^t \log \lambda_k(\theta(s-), X(s-), s-) dY_k(s) - \int_0^t \Big[\lambda_k - 1 \Big] ds \right\}.$$

Define: $\phi(f,t) = E^Q[f(\theta(t), X(t))L(t)|\mathcal{F}_t^{\vec{Y}}]$.
Then, $\phi(1,t)$ is the likelihood of Y or the **integrated (marginal) likelihood** of Y after assigning a prior to $(\theta(0), X(0))$.

Define: π_t is the conditional distribution of $(\theta(t), X(t))$ given $\mathcal{F}_t^{\vec{Y}}$.
π_t becomes the **posterior** after a prior is assigned.

Define: $\pi(f,t) = E^P[f(\theta(t), X(t))|\mathcal{F}_t^{\vec{Y}}] = \int f(\theta, x)\pi_t(d\theta, dx)$.

Kallianpur-Striebel (Bayes) Formula gives:

$$\pi(f,t) = \frac{\phi(f,t)}{\phi(1,t)}.$$

Bayes Factor

Suppose there are two models: Model 1 and Model 2. **Define:**

$$q_1(f_1, t) = \frac{\phi_1(f_1, t)}{\phi_2(1, t)} \quad \text{and} \quad q_2(f_2, t) = \frac{\phi_2(f_2, t)}{\phi_1(1, t)}$$

The Bayes Factors: (the ratio of two integrated likelihoods)

$$BF_{12} = \frac{\phi_1(1, t)}{\phi_2(1, t)} = q_1(1, t) \quad \text{and} \quad BF_{21} = \frac{\phi_2(1, t)}{\phi_1(1, t)} = q_2(1, t)$$

- **Strongly Reject Model 1** if BF_{21} is larger than 12.
- **Decisively Reject Model 1** if BF_{21} is larger than 150.

Advantage: (1) Bayes factors do not require the two models to be nested, nor their distributions to be absolutely continuous w.r.t. each other.
(2) Under some conditions, *Bayes factor* $\approx$ BIC, which penalizes according to both the number of parameters and the number of data.

Filtering and Evolution Equations

Theorem 1.1: *Under Assumptions 1–5,*

$$\phi(f,t) = \phi(f,0) + \int_0^t \phi(\mathbf{A}f - (a-n)f, s)ds + \sum_{j=1}^{n} \int_0^t \phi((ap_j - 1)f, s-)dY_j(s).$$

Assuming $a_k(\theta(t), X(t), t) = a(t)$ *for* $\pi(f,t), q_k(f,t)$, $k = 1, 2$,

$$\pi(f,t) = \pi(f,0) + \int_0^t \pi(\mathbf{A}f, s)ds + \sum_{j=1}^{n} \int_0^t \Big[\frac{\pi(fp_j, s-)}{\pi(p_j, s-)} - \pi(f, s-)\Big] dY_j(s),$$

Assume Model 1 has generator, $\mathbf{A_1}$ *and Model 2 has generator,* $\mathbf{A_2}$.

$$q_1(f_1,t) = q_1(f_1,0) + \int_0^t q_1(\mathbf{A_1}f_1, s)ds + \sum_{j=1}^{n} \int_0^t \Big[\frac{q_1(f_1 p_j^{(1)}, s-)}{q_2(p_j^{(2)}, s-)} q_2(1, s-) - q_1(f_1, s-)\Big] dY_j(s)$$

$q_2(f_2, t) = ...$

where $p_j^{(k)} = p(y_j|x)$ *is the transition probability from* x *to* y_j *in Model* k, $k = 1, 2$.

A Limit Theorem and Bayesian Inference via Filtering

Theorem 1.2: *Suppose that Assumptions 1 to 5 hold for* $(\theta, X, \vec{Y})$ *and* $(\theta_\epsilon, X_{\epsilon_x}, \vec{Y}_\varepsilon)$. *If* $(\theta_\epsilon, X_{\epsilon_x}) \Rightarrow (\theta, X)$ *as* $\varepsilon = \max\{\varepsilon_x, |\epsilon|\} \to 0$, *then for bounded continuous functions,* f, f_1 *and* f_2,
(i) $\vec{Y}_\varepsilon \Rightarrow \vec{Y}$, *(ii)* $\phi_\varepsilon(f,t) \Rightarrow \phi(f,t)$, *(iii)* $\pi_\varepsilon(f,t) \Rightarrow \pi(f,t)$
and (iv) $q_{k,\varepsilon}(f_k,t) \Rightarrow q_k(f_k,t)$ *for* $k = 1, 2$ *simultaneously.*

Three-Step Construction of Recursive Algorithms – by Markov chain approximation method *for computing posterior, integrated likelihood and Bayes factors*

- Construct a continuous-time Markov chain $(\theta_\epsilon, X_\epsilon)$ to approximate (θ, X).
- Derive the filtering (or evolution) equations for $(\theta_\epsilon, X_\epsilon, \vec{Y}_\varepsilon)$.
- Convert the equation for $(\theta_\epsilon, X_\epsilon, \vec{Y}_\varepsilon)$ to recursive algorithms by

(a) representing $\pi_\epsilon(\cdot, t)$, for example, as a finite array with components being $\pi_\epsilon(f,t)$ for lattice-point indicator f;
(b) approximating the time integral with an Eular scheme.

The Proposed Framework

I. Filtering with a MPP observation

• **Setup:** measurable *mark space:* $(U, \mathcal{U}, \mu)$, μ: σ-finite;
$\mathcal{A} = \{A \in \mathcal{U} : \mu(A) < \infty\}$.; N is a Poisson Random Measure (PRM) on $\mathcal{U} \times \mathcal{B}[0, \infty)$ with mean measure $\mu \times m$.

UHF data can be viewed as an observed sample path of the MPP, Y,

$$Y(A, t) = Y(A \times [0, t]) = N\Big(\int_0^t \int_A \lambda(\theta(s), X(s), V(s); u, s)\mu(du)ds\Big), \quad \text{for} A \in \mathcal{A}$$

where $Y(A, t)$ is a counting process recording the cumulative number of events that have occurred in the set A up to time t.

Assumption 2.1: is the same as Assumption 1.1, but both θ and X are vector processes. Or, (θ, X) is a semimartingale vector processes.
Assumption 2.2: N is a PRM with mean measure $\mu \times m$ under P.
Assumption 2.3: (θ, X) and N are independent under measure P.
Assumption 2.4: $0 \leq a(\theta, x, v, t) \leq C$ for for some $C > 0$ and all $t > 0$.
Assumption 2.5: Intensity kernel: $\lambda(\theta, x, v; u, t) = a(\theta, x, v, t)p(u|x; \theta, v)$, where $p(u|x; \theta, v)$ is the transition probability from x to u.

II. Construction of the Partially-Observed Model

Value: X; Price: Y; Parameters in (X, Y): $\theta(t)$.

- Specify jointly (θ, X) as in Assumption 2.1.

- Trading times are driven by a conditional Poisson process with the intensity $a(\theta(t), X(t), V(t), t)$ with Assumption 2.4 and the trading times are $t_1, t_2, \ldots, t_i, \ldots$.

- The price is modeled as

$$Y(t_i) = F(X(t_i)),$$

where $F(\cdot)$ is a random transformation with a transition probability $p(y|x; \theta, v)$.

Remark: The two models are equivalent in distribution.

Examples

- Many models in Direction Three under the two Views such as Exponential ACD model, UHF-GARCH and more.

- Estimating Volatility via filtering: Frey and Runggaldier (2001) and Cvitanic, Liptser and Rozovskii (2003).

- Estimating Markov process sampled at conditional Poisson time: Duffie and Glenn (2003).

- Classical examples of MPP filtering problems: Chapter 6.3 of Bremaud (1981), Chapter 19.3 of Liptser and Shiryayev (2002) and Chapter 11 of Last and Brandt (1995).

Likelihoods

The Joint Likelihood: Assume μ is a finite measure. Then,

$$L(t) = \frac{dP}{dQ}(t) = \exp\left\{ \int_U \int_0^t \log \lambda(\theta(s-), X(s-), V(s-); u, s-) Y(du \times ds) \right.$$
$$\left. - \int_U \int_0^t \Big[\lambda(\theta(s), X(s), V(s); u, s) - 1 \Big] \mu(du) ds \right\}$$

Let $\phi(f,t) = E^Q[f(\theta(t), X(t))L(t)|\mathcal{F}_t^{Y,\vec{V}}]$. The unnormalized filtering equation:

$$\phi(f,t) = \phi(f,0) + \int_0^t \phi(\mathbf{A}f, s)ds - \int_U \int_0^t \phi((\lambda - 1)f, s)\mu(du)ds$$
$$+ \int_U \int_0^t \phi((\lambda - 1)f, s-)Y(du \times ds),$$

where $\lambda = \lambda(\theta(s), X(s), V(s); u, s)$.

Future Works

1. **More Statistical Foundations**
   (a) Continuous-time Posterior, and Likelihood Ratio or Bayes Factors
   (b) Filtering Equations and Evolution Equations for Bayes Factors
2. **Mathematical Foundations for Computational Algorithms**
   (a) Limit theorems and the Markov chain approximation method
   (b) Limit theorems and particle filtering or sequential Monto Carlo
3. **Applications to the econometrics of UHF data**
   (a) multi-asset case of Zeng (2003) allowing correlations among assets, stochastic volatilities and trading noises.
   (b) extend the dynamics of discrete bid and ask quotes of Hasbrouck (1999) with random arrival times and with clustering and non-clustering noises.
4. Option price and Hedge Portfolio for models which incorporate "transient" noises, trade-by-trade.
5. Asymptotic properties of the parameter estimators.