Maximum Likelihood Estimation for Multivariate Diffusions via an Itô-Taylor Expansion Approach

Xiangwei Wan*

(A Joint Work with Nan Chen† and Nian Yang ‡)

*Shanghai Jiao Tong University

†The Chinese University of Hong Kong

‡Nanjing University

The Fifth Asian Quantitative Finance Conference
April 25,2017, Seoul, Korea

Motivation

- Continuous-time models are widely used in financial economics for modeling the dynamics of asset returns, interest rates
- In financial econometrics, parameter estimation of continuous-time models based on discretely observed data is a key theme and maximum-likelihood estimation is the method of choice because of its consistency and efficiency.
- However, for most models, likelihood functions (functional of transition density function) are analytically intractable and thus involve heavy computational load in the optimization process for the maximum-likelihood estimators (MLE) .

Related Literatures: Closed-form Approximation

- Aït-Sahalia(02, Econometrica): univariate diffusions, Hermite series expansion
  - Lamperti transform: reducible and irreducible diffusions
  - Refinements: Bakshi and Ju(05, J. Bus.), Lee et al (14, JoE)
  - Extensions: Time-inhomogeneous (Egorov et al, 03, JoE) ; Jump-diffusion (Schaumburg, 01);
- Aït-Sahalia(08, Ann. Stat.): multivariate diffusions, Taylôr expansion in (small) time and states to match Kolmogorov equation
  - Extensions: Time-inhomogeneous (Choi, 13, 15, JoE) ; Jump-diffusion (Yu, 07, JoE)
- Filipović et al(13, JoE): multivariate affine jump-diffusion, orthonormal polynomials (w.r.t an auxiliary density) series expansion
- Li (13, Ann. Stat.), Li and Chen (16, JoE): Time homogeneous multivariate jump-diffusions, pathwise small time Taylôr expansion

Itô-Taylor Expansion

- Let $f(t,x)$ be $J+1$ and $2(J+1)$ continuous differentiable in t and in x, respectively, applying Itô formula repeatedly

$$\mathbf{E}[f(t+\Delta, X(t+\Delta))|X(t)=x] = \sum_{n=0}^{J} \frac{\Delta^n}{n!}(\partial_t + \mathcal{L}_x)^n f(t,x) + \mathcal{O}(\Delta^{J+1}),$$

where $\mathcal{L}_x$ is the infinitesimal generator of the multivariate time-inhomogeneous diffusion $X(t)$

$$\mathcal{L}_x f(t,x) = \sum_{i=1}^{m} \mu_i(t,x)\frac{\partial f(t,x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{m} \nu_{i,j}(t,x)\frac{\partial^2 f(t,x)}{\partial x_i \partial x_j}.$$

- The transition density p of diffusion $X(t)$ can be written as a conditional expectation form (Watanabe,1987), that is,

$$p(t', x'|t, x; \theta) = \mathbf{E}[\delta(X(t') - x')|X(t) = x].$$

- But, $\delta(\cdot)$ is the Dirac Delta function, which is not smooth.

Offsetting Non-Smoothness and the New Problem

- Use normal distribution to approximate Dirac Delta function.
- Define $q(t', x'; t, x)$ as the transition density of the diffusion with zero drift and fixed variance-covariance matrix ν_0. Then,

$$\begin{aligned} p(t', x'|t, x; \theta) &= \mathbf{E}[\delta(X(t') - x')|X(t) = x] \\ &= \lim_{s \uparrow t'} \mathbf{E}[q(t', x'; s, X(s))|X(t) = x] \\ &= \sum_{i=0}^{J} \frac{(t'-t)^i}{i!} (\partial_t + \mathcal{L}_x)^i q(t', x'; t, x) + \mathcal{R}_J, \end{aligned}$$

- New Problem: $q(t', x'; t, x)$ is a function of $(x' - x)/\sqrt{t' - t}$, and thus its $2i$-th derivative on x generates $(t' - t)^{-i}$. So, $\mathcal{R}_J$ might be $\mathcal{O}(1)$, and the expansion may not converge as $t' - t$ approaches zero!
- To overcome the problem, we first give a decomposition:

$$(\partial_t + \mathcal{L}_x)^i q(t', x'|t, x) = \sum_{|h|=1}^{2i} w_{i,h}(t, x) \cdot \partial_h q(t', x'|t, x)$$

Our Contribution

Lemma

Fix ν_0 on the initial variance-covariance, i.e. $\nu_0 = \nu(t, x)$, we have

$$w_{i,h}(t, x) = 0, \text{ for all } |h| > 3i/2.$$

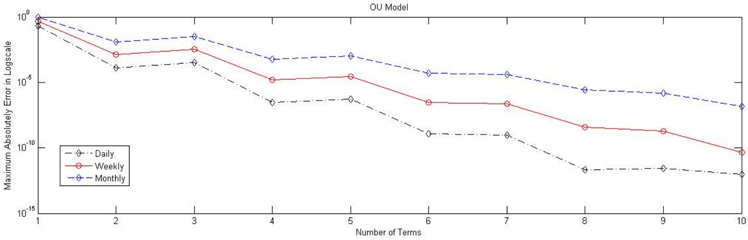
- Thus, the derivatives $\partial_h q(t', x'|t, x)$ will offset at most $(t' - t)^{\frac{3i}{4}}$, and the general term of our special Itô-Taylor expansion is of order $(t' - t)^{\frac{i}{4} - \frac{m}{2}}$.

Theorem

Fixed ν_0 as above, we have the Itô-Taylor expansion:

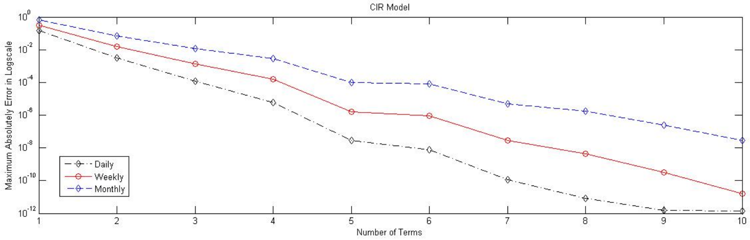
$$\begin{aligned} p(t', x'|t, x; \theta) &= \sum_{i=0}^{J} \frac{(t'-t)^i}{i!} (\partial_t + \mathcal{L}_x)^i q(t', x'; t, x) + \mathcal{O}((t'-t)^{\frac{J+1}{4} - \frac{m}{2}}) \\ &= \sum_{i=0}^{J} \frac{(t'-t)^i}{i!} \sum_{|h| \leq 3i/2} w_{i,h}(t, x) \partial_h q(t', x'|t, x) + \mathcal{O}((t'-t)^{\frac{J+1}{4} - \frac{m}{2}}) \end{aligned}$$

Absolute error of approximated density for OU process



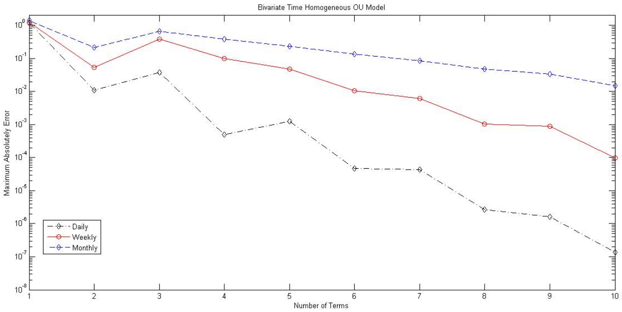
- The absolute error of OU process between this paper's approximated density and the true transition density
- OU process: $dX_t = 0.5(0.06 - X_t)dt + 0.03dW_t$.
- Monitored frequency: monthly (blue), weekly (red), and daily (black) .

Absolute error of approximated density for CIR process



- The absolute error of CIR process between this paper's approximated density and the true transition density
- CIR process: $dX(t) = 0.5(0.06 - X(t))dt + 0.15\sqrt{X(t)}dW(t)$.
- Monitored frequency: monthly (blue), weekly (red), and daily (black).

Absolute error of approximated density for BOU process

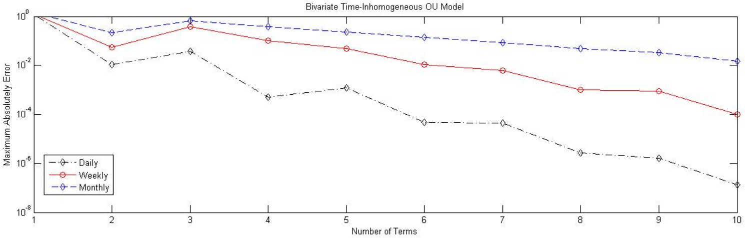


- The absolute error of BOU process between this paper's approximated density and the true transition density
- Bivariate Ornstein-Uhlenbeck (BOU) process:

$$dX(t) = \begin{pmatrix} 5 & 0 \\ 1 & 10 \end{pmatrix} \begin{pmatrix} 0 - X_1(t) \\ 0 - X_2(t) \end{pmatrix} dt + dW(t),$$

- Monitored frequency: monthly (blue), weekly (red), and daily (black).

Absolute error of approximated density for BOUI process

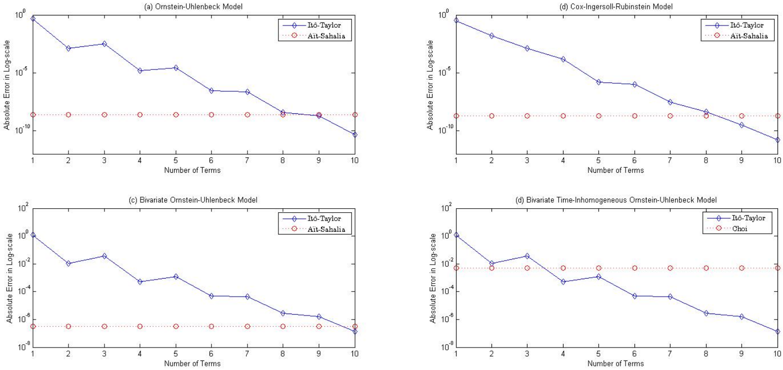


- The absolute error of BOUI process between this paper's approximated density and the true transition density
- Bivariate Time-inhomogeneous Ornstein-Uhlenbeck (BOUI) process:

$$dX(t) = \begin{pmatrix} 5 & 0 \\ 1 & 10 \end{pmatrix} \begin{pmatrix} 0 + 0.1t - X_1(t) \\ 0 + 0.1t - X_2(t) \end{pmatrix} dt + dW(t),$$

- Monitored frequency: monthly (blue), weekly (red), and daily (black).

Comparison of Maximum Absolute Error between Itô-Taylor Expansion and Aït-Sahalia (2008), Choi(2013)



Comparison of Maximum Absolute Error between our Itô-Taylor Expansion and a third-order expansion of Aït-Sahalia (2008), and a second-order expansion of Choi(2013).

The Convergence of Our Approximate MLE

- θ_0: the true value of the parameter vector
- $\hat{\theta}_n$: MLE using the true transition density with n samples
- $\hat{\theta}_n^{(J)}$: MLE using the J-th order approximate transition density

Theorem

Fix the sample size n and $J > 2m - 1$. We have

$$\hat{\theta}_n^{(J)} - \hat{\theta}_n \to 0$$

in $\mathbb{P}_{\theta_0}$-probability as $t' - t \to 0$.

Approximate MLE for Bivariate OU process via Monte Carlo

θ_0	$\hat{\theta}_n - \theta_0$	$\hat{\theta}_n^{(AS)} - \hat{\theta}_n$	$\hat{\theta}_n^{(4)} - \hat{\theta}_n$	$\hat{\theta}_n^{(5)} - \hat{\theta}_n$	$\hat{\theta}_n^{(6)} - \hat{\theta}_n$	$\hat{\theta}_n^{(7)} - \hat{\theta}_n$
$\kappa_{11} = 5$	0.41888	0.00186	0.00774	-0.00368	0.00222	0.00080
	(1.11974)	(0.04472)	(0.14139)	(0.10338)	(0.06851)	(0.06916)
$\kappa_{21} = 1$	0.03538	0.00105	0.01423	-0.00017	0.00276	-0.00123
	(1.19813)	(0.02933)	(0.15667)	(0.11368)	(0.08693)	(0.05589)
$\kappa_{22} = 10$	0.61443	0.00324	0.04357	-0.000490	0.00239	-0.00097
	(1.54834)	(0.02971)	(0.27662)	(0.22413)	(0.10551)	(0.10502)
$\alpha_1 = 0$	0.00486	0.00091	0.00021	0.00105	0.00132	0.00052
	(0.06284)	(0.01477)	(0.02000)	(0.01959)	(0.01962)	(0.01983)
$\alpha_2 = 0$	-0.00033	0.00009	0.00003	-0.00008	0.00003	-0.00003
	(0.03351)	(0.00443)	(0.00675)	(0.00653)	(0.00637)	(0.00563)

Notes: We use θ_0 to generate 1000 sample paths. Each of them contains 500 weekly observations (i.e., we take $t' - t = 1/52$). The first column reports true parameter values θ_0. The second column reports the bias and the standard derivation (values in parentheses) of the true maximum likelihood estimator $\hat{\theta}_n$. The third column shows the difference between true maximum likelihood estimator $\hat{\theta}_n$ and the 3rd-order approximate estimator $\hat{\theta}_n^{(AS)}$ developed by Aït-Sahalia (2008). The remaining columns report the differences between true maximum likelihood estimator $\hat{\theta}_n$ and Itô-Taylor expansion estimator $\hat{\theta}_n^{(J)}$ developed in this paper, with the standard derivation in parentheses. The order of our Itô-Taylor expansion takes values from $J = 4$ to $J = 7$.

Approximate MLE for Bivariate Time-Inhomogeneous OU process via Monte Carlo

Table: Monte Carlo Evidence for BOUI Model

θ_0	$\hat{\theta}_n - \theta_0$	$\hat{\theta}_n^{(Choi)} - \hat{\theta}_n$	$\hat{\theta}_n^{(4)} - \hat{\theta}_n$
$\kappa_{11} = 5$	0.59446	0.08046	-0.01367
	(1.17004)	(0.26004)	(0.33544)
$\kappa_{21} = 1$	0.10497	-1.11154	0.00273
	(1.24774)	(1.94661)	(0.34327)
$\kappa_{22} = 10$	0.62750	-0.37080	-0.11460
	(1.57026)	(0.38605)	(0.52059)
$\alpha_1 = 0$	-0.00163	0.00546	0.00263
	(0.08942)	(0.07228)	(0.07931)
$\alpha_2 = 0$	0.00205	0.00137	-0.00024
	(0.06047)	(0.03442)	(0.03417)
$\beta_1 = 0.1$	0.00112	-0.00057	-0.00047
	(0.01839)	(0.01124)	(0.01241)
$\beta_2 = 0.1$	-0.00085	0.00046	0.00007
	(0.01122)	(0.00540)	(0.00560)

Notes: We use θ_0 to generate 1000 sample paths. Each of them contains 500 weekly observations .

Additional Numerical Results

- Hestion model (Affine model)

$$\begin{cases} dX_t & = (\mu - \frac{1}{2}X_t)dt + \sqrt{V_t}X_t[\rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}] \\ dV_t & = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_{1,t} \end{cases}$$

where X_t is the underlying asset return and V_t is its volatility,

- Continuous-time GARCH model (Non-affine)

$$\begin{cases} dX_t & = (\mu - \frac{1}{2}X_t)dt + \sqrt{V_t}X_t[\rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}] \\ dV_t & = \kappa(\theta - V_t)dt + \sigma V_t dW_{1,t} \end{cases}$$

Approximate MLE for Heston and GARCH model via Monte Carlo Simulation

Table: Monte Carlo Evidence for Heston and GARCH Models

	Heston Model ($\beta = 1/2$)		GARCH Model ($\beta = 1$)	
θ_0	$\hat{\theta}_n^{(AS)} - \theta_0$	$\hat{\theta}_n^{(4)} - \theta_0$	$\hat{\theta}_n^{(AS)} - \theta_0$	$\hat{\theta}_n^{(4)} - \theta_0$
$\sigma = 0.25$	0.00009	0.00008	0.00060	0.00060
	(0.00625)	(0.00625)	(0.00645)	(0.00649)
$\rho = -0.8$	-0.00071	-0.00071	-0.00069	-0.00068
	(0.01294)	(0.01291)	(0.01338)	(0.01339)
$\alpha = 0.1$	0.00496	0.00267	-0.00003	0.00002
	(0.03783)	(0.02901)	(0.00225)	(0.00228)
$\mu = 0.03$	-0.02700	-0.02855	0.00219	0.00013
	(0.19754)	(0.19767)	(0.07899)	(0.07998)
$\kappa = 3$	0.76111	0.77456	0.15436	0.15563
	(1.52510)	(1.51853)	(0.53960)	(0.54548)

Notes: We use θ_0 to generate 1000 sample paths. Each of them contains 500 weekly observations.

Conclusion

- Itô-Taylor expansion is simple to implement, only the differentiation is needed
- We prove that our special Itô-Taylor expansion is convergent
- The method nests multivariate time homogeneous and time-inhomogeneous diffusions, reducible and irreducible diffusions, affine and non-affine diffusions.
- Monte Carlo experiments show that this method is quite accurate and numerically stable.

Thanks and Comments!