Topic 9: Continuous Time Models

Jun Yu

School of Economics, Singapore Management University

March 11, 2010

Why Model in Continuous Time

- The economy and financial markets are continuously operating. The underlying decision processes involves millions of decisions.
- There are two types of variables: stock variables and flow variables. Continuous time methods allow for the correct treatment of them.
- Observed frequency may be different from forecasting frequency. Variables may be observed at different frequencies.
- Continuous systems allow for more realistic specifications of the partial adjustment processes.
- Natural applications in the development of finance theory and derivative pricing models.

Ordinary Differential Equations: Theory

- General differential equation with an initial condition (i.e. the initial value problem – IVP)

$$\frac{dY(t)}{dt} = f(t, Y), Y(0) = Y_0 \tag{1}$$

- In general no analytical solution to Equation (1) is available. For some special f, the analytical solution is available.
- Example 1:

$$\frac{dY(t)}{dt} = -\kappa Y, Y(0) = Y_0 \tag{2}$$

$$\frac{dY(t)}{Y(t)} = -\kappa dt \Rightarrow \frac{d \ln Y}{dt} = -\kappa$$

$$\ln Y(t) = -\kappa t + \ln Y_0 \Rightarrow Y(t) = Y_0 e^{-\kappa t} \tag{3}$$

Ordinary Differential Equations: Theory

- Example 2:

$$\frac{dY(t)}{dt} = \kappa(\mu - Y), Y(0) = Y_0 \tag{4}$$

$$\frac{d(Y(t) - \mu)}{Y(t) - \mu} = -\kappa dt \Rightarrow \frac{d \ln X}{dt} = -\kappa$$

$$\ln X(t) = -\kappa t + \ln X_0 \Rightarrow Y(t) = \mu + (Y_0 - \mu)e^{-\kappa t} \tag{5}$$

Ordinary Differential Equations: Theory

- Example 3:

$$\frac{dY(t)}{dt} = -\kappa Y + b(t) = p(t)Y + b(t), \, Y(0) = Y_0 \tag{6}$$

where $p(t) = -\kappa$ but it can be more general.

- Define the integration factor $v(t) = \exp\left(-\int_0^t p(r)dr\right) = \exp(\kappa t)$
- By definition, $dv(t)/dt = -v(t)p(t) = \kappa v(t)$. Consider

$$\frac{d(v(t)Y(t))}{dt} = v(t)\frac{dY(t)}{dt} + Y(t)\frac{dv(t)}{dt} = v(t)b(t)$$

- The solution is

$$v(t)Y(t) - v(0)Y(0) = \int_0^t v(r)b(r)dr = \int_0^t \exp(\kappa r)b(r)dr$$

$$Y(t) = Y_0 \exp(-\kappa t) + \int_0^t \exp(-\kappa(t-r))b(r)dr \tag{7}$$

Ordinary Differential Equations: Theory

- Example 4:

$$\frac{dY(t)}{dt} = \kappa(\mu - Y) + b(t), Y(0) = Y_0 \tag{8}$$

- The solution is

$$Y(t) = \mu + (Y_0 - \mu)\exp(-\kappa t) + \int_0^t \exp(-\kappa(t-r))b(r)dr \tag{9}$$

Ordinary Differential Equations: Theory

- **Picard–Lindelöf Theorem**: For the following IVP

  $$\frac{dY(t)}{dt} = f(t, Y), Y(0) = Y_0$$

  suppose f is Lipschitz continuous in y and continuous in t. Then, there exists a unique solution $Y(t)$ to the IVP.
- **Lipschitz continuous**: $f(y)$ is Lipschitz continuous with a Lipschitz constant L iff for all y_1, y_2 such that $|f(y_1) - f(y_2)| \leq L|y_1 - y_2|$.

Ordinary Differential Equations: Numerical Issues

- An IVP is stable if there exists C and η such that for all $\varepsilon \in (0, \eta)$, when $|\varepsilon_0| < \varepsilon, |\delta(t)| < \varepsilon$, the solution of the perturbed IVP

$$\frac{dZ(t)}{dt} = f(t, Z) + \delta(t), Z(0) = Y_0 + \varepsilon_0 \tag{10}$$

satisfies $|Y(t) - Z(t)| < C\varepsilon$.

Ordinary Differential Equations: Numerical Issues

- A numerical solution is unstable if errors grow exponentially for a problem for which there is a bounded solution.
- Stability can depend on three factors: IVP, numerical method and step size.

Stochastic Processes and Stochastic Differential Equations

- **Euler method**. Split the interval $[0, T]$ into T/h subintervals so that each interval has a size of h (known as the step size), i.e., $t_i = 0, 1/h, 2/h, ...$ for $i = 0, 1, ...$ Replace the differentiation by the difference,

$$\begin{aligned} \frac{Y_i - Y_{i-1}}{h} &\approx f(t_{i-1}, Y_{i-1}), Y(0) = Y_0, \\ Y_i &= Y_{i-1} + hf(t_{i-1}, Y_{i-1}), i = 1, ..., T/h. \end{aligned}$$

  Euler method approximates $Y(t_i)$ by Y_i obtained from the above recursions.

- Local truncation error: $O(h^2)$. Global truncation error: $O(h)$.

Stochastic Processes and Stochastic Differential Equations

- **Trapezoidal rule**:

$$\frac{Y_i - Y_{i-1}}{h} = 0.5(f(t_{i-1}, Y_{i-1}) + f(t_i, Y_i)), Y(0) = Y_0.$$

  This is implicit.

- Local truncation error: $O(h^3)$. Global truncation error: $O(h^2)$.
- **Heun's method**:

$$Y_i^0 = Y_{i-1}^m + hf(t_{i-1}, Y_{i-1})$$

$$Y_i^k = Y_{i-1} + 0.5h\left(f(t_{i-1}, Y_{i-1}^m) + f(t_i, Y_i^{k-1})\right), k = 1, ..., m$$

  This is explicit.

- Local truncation error: $O(h^3)$. Global truncation error: $O(h^2)$.

Stochastic Processes and Stochastic Differential Equations

- **The midpoint method (second order Runge-Kutta method)**:

$$Y_{i+1/2} = Y_i + 0.5hf(t_i, Y_i)$$

$$Y_{i+1} = Y_i + hf(t_{i+1/2}, Y_{i+1/2})$$

- Local truncation error: $O(h^3)$. Global truncation error: $O(h^2)$.

Stochastic Processes and Stochastic Differential Equations

- **The general Runge-Kutta method**:

$$Y_{i+1} = Y_i + h\phi$$

$$\phi = a_1 k_1 + a_2 k_2 + \cdots + a_n k_n$$

$$\begin{aligned} k_1 &= f(t_i, y_i), k_2 = f(t_i + p_1 h, y_i + q_{11} k_1 h), \cdots \\ k_n &= f(t_i + p_{n-1} h, y_i + q_{n-1,1} k_1 h + \cdots + q_{n-1,n-1} k_{n-1} h) \end{aligned}$$

- Coefficients $a_i, p_i, q_{i,j}$ are obtained by matching the difference equation with the Taylor series.

Stochastic Processes and Stochastic Differential Equations

- A Brownian motion has no drift and has a variance rate of 1 (i.e., the variance over one unit time interval is 1). A generalized Wiener process allows for a variance rate of σ^2, and is defined by

  $$dX(t) = \sigma\, dB(t).$$

  The above equation can be equivalently written as $X(t) = \int_0^t \sigma\, dB(r)$ where

  $$\int_0^t H\, dB(r) = \lim_{n\to\infty} \sum_{t_i, t_{i-1} \in [0,t]} H_{t_{i-1}}(B(t_i) - B(t_{i-1})).$$

Stochastic Processes and Stochastic Differential Equations

- The generalized Wiener process which allows for a *drift rate* of a and a variance rate of σ^2 is defined by

$$dX(t) = a\,dt + \sigma\,dB(t)$$

- The transition density is

$$X(t+h)|X(t) \sim N(ah + X(t), \sigma^2 h).$$

Stochastic Processes and Stochastic Differential Equations

- Stochastic differential equation based on the Brownian motion (i.e. diffusion):

$$dY(t) = \mu(t, Y)dt + \sigma(t, Y)dB(t), Y(0) = Y_0 \tag{11}$$

- In general no analytical solution to Equation (11) is available. For some special μ and σ, the analytical solution is available.
- Example 1:

$$dY(t) = dB(t), Y(0) = 0 \tag{12}$$

$$Y(t) = \int_0^t dB(r) = B(t)$$

Stochastic Processes and Stochastic Differential Equations

- Example 2:

$$dY(t) = \mu dt + \sigma dB(t), Y(0) = Y_0 \tag{13}$$

$$Y(t) = \int_0^t \mu dt + \int_0^t \sigma dB(r) = \mu t + \sigma B(t)$$

- The transition density is

$$X(t+h)|X(t) \sim N(ah + X(t), \sigma^2 h).$$

Stochastic Processes and Stochastic Differential Equations

- Example 3 (OU):

$$dY(t) = -\kappa Y dt + dB(t), Y(0) = 0 \tag{14}$$

Similar to (7),

$$Y(t) = \int_0^t \exp(-\kappa(t-r)) dB(r) := J_{-\kappa}(t) \tag{15}$$

So for any h

$$Y(t+h) = e^{-\kappa h} Y(t) + \int_0^h \sigma \exp(-\kappa(h-s)) dB(t+s) \tag{16}$$

Stochastic Processes and Stochastic Differential Equations

- Example 4 (OU):

$$dY(t) = -\kappa Y dt + \sigma dB(t), Y(0) = Y_0 \tag{17}$$

Similar to (7),

$$Y(t) = Y_0 e^{-\kappa t} + \int_0^t \sigma \exp(-\kappa(t-r)) dB(r) = Y_0 e^{-\kappa t} + \sigma J_{-\kappa}(t) \tag{18}$$

So for any h

$$Y(t+h) = e^{-\kappa h} Y(t) + \int_0^h \sigma \exp(-\kappa(h-s)) dB(t+s) \tag{19}$$

Stochastic Processes and Stochastic Differential Equations

- This is an AR(1) model. Since

$$\begin{aligned}\int_0^h \sigma e^{-\kappa(h-s)} dB(t+s) &\sim N\left(0, \int_0^h \sigma^2 e^{-2\kappa(h-s)} ds\right)\\ &= N\left(0, \frac{\sigma^2(1-\exp(-2\kappa h))}{2\kappa}\right),\end{aligned}$$

  the transition density is

$$Y(t+h)|Y(t) \sim N\left(e^{-\kappa h} Y(t), \frac{\sigma^2(1-\exp(-2\kappa h))}{2\kappa}\right). \quad (20)$$

Stochastic Processes and Stochastic Differential Equations

- Example 5 (Vasicek):

  $$dY(t) = \kappa(\mu - Y)dt + \sigma dB(t), Y(0) = Y_0 \tag{21}$$

  Similar to (9),

  $$Y(t) = \mu + (Y_0 - \mu)\exp(-\kappa t) + \int_0^t \sigma \exp(-\kappa(t-r))dB(r)$$

- μ is the mean and $\kappa(>0)$ is the speed of mean reversion.

Stochastic Processes and Stochastic Differential Equations

- For any h

$$Y(t+h) = \mu + e^{-\kappa h}(Y(t) - \mu) + \int_0^h \sigma \exp(-\kappa(h-s))dB(t+s)$$

- The transition density is

$$Y(t+h)|Y(t) \sim N\left(\mu + (Y_0 - \mu)\exp(-\kappa h), \frac{\sigma^2(1 - e^{-2\kappa h})}{2\kappa}\right). \tag{22}$$

Stochastic Processes and Stochastic Differential Equations

- Example 6 (Variance Gamma Process):

$$dY(t) = \kappa(\mu - Y)dt + \sigma dL(t), Y(0) = Y_0 \qquad (23)$$

where $L(t) = B(\Gamma(t; 1/\nu, \nu))$ is the variance gamma process, also known as the Laplace motion and $\Gamma(\cdot; c, \nu)$ is a gamma distribution with mean parameter c and variance parameter ν.

- Similar to (16),

$$Y(t+h) = \mu + e^{-\kappa h}(Y(t) - \mu) + \sigma\sqrt{\frac{1 - e^{-2\kappa h}}{2\kappa}}\varepsilon_t \qquad (24)$$

where ε_t follows the variance gamma distribution.

Stochastic Processes and Stochastic Differential Equations

- Example 7 (CIR):

$$dY(t) = \kappa(\mu - Y)dt + \sigma\sqrt{Y}dB(t), Y(0) = Y_0 \tag{25}$$

- Consequently,

$$Y(t) = [Y(0) - \mu]e^{-\kappa t} + \mu + \sigma \int_0^t e^{-\kappa(t-s)}\sqrt{Y}dB(s)$$

$$\begin{aligned} Y(t+h) &= \mu\left(e^{-\kappa h} - 1\right) + e^{-\kappa h}Y(t) \\ &\quad + e^{-\kappa h}\int_0^h e^{-\kappa s}\sqrt{Y(t+s)}dB(t+s). \end{aligned}$$

Stochastic Processes and Stochastic Differential Equations

- Feller (1951) and CIR (1985) show that the transition density is

  $$pdf(Y(t+h)|Y(t)) = ce^{-u-v}(v/u)^{q/2} I_q(2(uv)^{1/2})$$

  and the marginal density of $Y(t)$ is

  $$w_1^{w_2} r^{w_2-1} e^{-w_1 r}/\Gamma(w_2)$$

  where $c = 2\kappa/(\sigma^2(1-e^{-\kappa h}))$, $u = cY(t)e^{-\kappa h}$, $v = cY(t+h)$, $q = 2\kappa\mu/\sigma^2 - 1$, $w_1 = 2\kappa/\sigma^2$, $w_2 = 2\kappa\mu/\sigma^2$, and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q.

Stochastic Processes and Stochastic Differential Equations

- Example 8 (CKLS):

$$dY(t) = \kappa(\mu - Y)dt + \sigma Y^{\gamma} dB(t), Y(0) = Y_0 \tag{26}$$

- Consequently,

$$Y(t) = [Y(0) - \mu]e^{-\kappa t} + \mu + \sigma \int_0^t e^{-\kappa(t-s)} Y^{\gamma}(s) dB(s)$$

$$\begin{aligned} Y(t+h) &= \mu\left(e^{-\kappa h} - 1\right) + e^{-\kappa h} Y\left(t\right) \\ &\quad + e^{-\kappa h} \int_0^h e^{-\kappa s} Y^{\gamma}\left(t+s\right) dB\left(t+s\right). \end{aligned}$$

Stochastic Processes and Stochastic Differential Equations

- Definition: Ito Process is defined as

$$dY(t) = \mu(Y,t)dt + \sigma(Y,t)dB(t), Y(0) = Y_0 \qquad (27)$$

- Ito's Lemma: Suppose $Y(t)$ is an Ito process, and $X(t) = f(t, Y_t)$, Then $X(t)$ is also an Ito process. Specifically, if

$$dY(t) = \mu(t, Y(t))\ dt + \sigma(t,\ Y(t))\, dB(t)$$

then

$$dX(t) = \left[\frac{\partial f}{\partial Y}\mu + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial Y^2}\sigma^2\right]\ dt + \frac{\partial f}{\partial Y}\sigma\, dB(t).$$

Stochastic Processes and Stochastic Differential Equations

- Example 9 (Geometric Brownian Motion):

$$dY(t) = \mu Y(t)\,dt + \sigma Y(t)\,dB(t). \tag{28}$$

- Let $X = \ln Y$ and apply the Ito's lemma, we get

$$\begin{aligned} d\ln Y(t) &= \left[\frac{1}{Y}\cdot\mu Y + \frac{1}{2}\left(-\frac{1}{Y^2}\right)\sigma^2 Y^2\right]dt + \frac{\sigma Y}{Y}\,dB(t) \\ &= \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma\,dB(t) \end{aligned}$$

Stochastic Processes and Stochastic Differential Equations

- Example 10 (Inverse CIR):

$$dY(t) = \kappa(\mu - Y)Ydt + \sigma Y^{1.5}(t)\, dB(t). \qquad (29)$$

- Let $X = 1/Y$ and apply the Ito's lemma, we get

$$\begin{aligned} dX(t) &= \left[-Y^{-2}\kappa(\mu - Y)Y + \frac{1}{2}\frac{2}{Y^3}\sigma^2 Y^3 \right] dt - \frac{\sigma Y^{1.5}}{Y^2}\, dB(t) \\ &= \kappa\mu \left(\frac{\kappa + \sigma^2}{\kappa\mu} - X(t) \right) dt + \sigma\sqrt{X(t)} dB(t) \end{aligned}$$

- This is a CIR model with mean $\frac{\kappa+\sigma^2}{\kappa\mu}$ and speed of mean reversion $\kappa\mu$.

Parametric methods

- The model

$$dX(t) = \mu(X(t);\theta)dt + \sigma(X(t);\theta)dB(t), \tag{30}$$

- Assume $X(t)$ is recorded discretely at points $(h, 2h, \cdots, Nh(:= T))$ in the time interval $[0, T]$, where h is the discrete interval of observation of $X(t)$ and T is the time span of the data. The full sequence of N observations is $\{X_h, X_{2h}, \cdots, X_{Nh}\}$. In finance $h = 1/12$ $(1/52$ or $1/252)$.
- First, assume that $X(t)$ is univariate.

Parametric methods

- It has been argued that when the model is correctly specified, the preferred choice of estimator and preferred basis for inference should be maximum likelihood (ML). References: Aït-Sahalia (2002) and Durham and Gallant (2002)
- Reasons:
    1. Generality and wide applicability
    2. Desirable asymptotic properties: consistency, asymptotic normality and asymptotic efficiency

Parametric methods

- Suppose $p(X_{ih}|X_{(i-1)h},\theta)$ is the transition probability density. The Markov property of model (30) implies the log-likelihood function for the discrete sample

$$\ell_{TD}(\theta)=\ln(p(X_{ih}|X_{(i-1)h},\theta)).$$

- MLE is consistent, asymptotically normally distributed and asymptotically efficient under the usual regularity conditions in (stationary) dynamic models (Hall and Heyde, 1980; Billingsley, 1961).
- In nonstationary, nonergodic cases, the limit theory is no longer asymptotically normal and there are several possibilities, including various unit root, local to unity, mildly explosive and explosive limit distributions.
- In some cases, nonstationary models can be transformed into stationary models

Parametric methods

- Unfortunately, only in rare cases, do the transition density $\ln(p(X_{ih}|X_{(i-1)h},\theta))$ have analytical expressions.
- $p(X_{ih}|X_{(i-1)h},\theta)$ satisfies the "forward" equation:

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial y^2},$$

or the "backward" equation

$$\frac{\partial p}{\partial s} = -\frac{1}{2}\frac{\partial^2 p}{\partial x^2}.$$

Special Cases

- **Geometric Brownian Motion**:

$$dX(t) = \mu X(t)\, dt + \sigma X(t)\, dB(t). \tag{31}$$

Black and Scholes (1973) used this model to develop option price formula. For this model,

$$\ln\left(X_{ih} | X_{(i-1)h}\right) \sim \mathrm{N}\left(\left(\mu - \frac{\sigma^2}{2}\right) h + \ln(X_{(i-1)h}), \sigma^2 h\right)$$

Special Cases

- **Ornstein-Uhlenbeck (OU) process (or Vasicek model)**:

$$dX(t) = \kappa(\mu - X(t))dt + \sigma\, dB(t). \tag{32}$$

Vasicek (1977) used this process to describe the movement of short term interest rates. The exact discrete model is

$$X_{ih} = e^{-\kappa h} X_{(i-1)h} + \mu \left(1 - e^{-\kappa h}\right) + \sigma \sqrt{(1 - e^{-2\kappa h})/(2\kappa)}\epsilon_i. \tag{33}$$

The transition density is

$$X_{ih} | X_{(i-1)h} \sim N\left(\mu(1 - e^{-\kappa h}) + e^{-\kappa h} X_{(i-1)h}, \sigma^2(1 - e^{-2\kappa h})/(2\kappa)\right). \tag{34}$$

Special Cases

- **Square-root (or CIR) model:**

$$dX(t) = \kappa(\mu - X(t))dt + \sigma\sqrt{X(t)}\,dB(t). \tag{35}$$

Cox, Ingersoll and Ross (1985) also used this process to describe movements in short term interest rates. The exact discrete model corresponding to (35) is given by

$$X_{ih} = e^{-\kappa h}X_{(i-1)h} + \mu\left(1 - e^{-\kappa h}\right) + \sigma\int_{(i-1)h}^{ih} e^{-\kappa(ih-s)}\sqrt{X(s)}dB(s). \tag{36}$$

When $2\kappa\mu/\sigma^2 \geq 1$, X is distributed over the positive half line. Feller (1951) showed that the transition density is non-central χ^2 and the marginal density is gamma.

Special Cases

- **Inverse square-root model:**

  $$dX(t) = \kappa(\mu - X(t))X(t)dt + \sigma X^{1.5}(t)\, dB(t). \qquad (37)$$

  Ahn and Gao (1999) again used this process to model short term interest rates. When $\kappa, \mu > 0$, X is distributed over the positive half line. The transition density is

  $$X_{ih}|X_{(i-1)h} = c^{-1}e^{-u-v}(v)^{q/2+2}u^{-q/2}I_q(2(uv)^{1/2}) \qquad (38)$$

  where $c = 2\kappa\mu/(\sigma^2(1-e^{-\kappa\mu h}))$, $u = ce^{-\kappa\mu h}/X_{(i-1)h}$,
  $v = c/X_{ih}$, $q = 2(\kappa+\sigma^2)/\sigma^2 - 1$.

Parametric methods

- The Euler scheme approximates a general diffusion process by the following discrete time model

$$X_{ih} = X_{(i-1)h} + \mu(X_{(i-1)h}, \theta)h + \sigma(X_{(i-1)h}, \theta)\sqrt{h}\epsilon_i. \tag{39}$$

- The transition density is

$$X_{ih}|X_{(i-1)h} \sim N\left(X_{(i-1)h} + \mu(X_{(i-1)h}, \theta)h, \sigma^2(X_{(i-1)h}, \theta)h\right).$$

- For the Vasicek model, the Euler discrete approximation is

$$X_{ih} = \kappa\mu h + (1 - \kappa h)X_{(i-1)h} + \sigma N(0, h). \tag{40}$$

- Consistent for σ when $h \to 0$ (in-fill asymptotics) or $T \to \infty$ (long-span asymptotics) but consistent for μ and κ when $T \to \infty$.

Parametric methods

- When h is small, the Euler scheme should provide a good approximation to the exact discrete time model. However, when h is large, the Euler approximation can be poor.
- Consider the case where $\kappa = 1$ and $h = 1/12$, in which case $e^{-\kappa h}$ is 0.92 whereas $1 - \kappa h$ is 0.9167 and the approximation is good. But if $\kappa = 1$ and $h = 1$, then $e^{-\kappa h}$ is 0.3679 whereas $1 - \kappa h$ is 0. These comparisons suggest that the Euler discretization offers a good approximation to the exact discrete time model for daily or higher frequencies but not for annual or lower frequencies.
- The bias introduced by this discrete time approximation is called the *discretization bias.*

Parametric methods

- A closely related discretization method, suggested by Bergstrom (1966), is based on integrating the stochastic differential equation and using the trapezoidal rule approximation

$$\int_{(i-1)h}^{ih} \mu(X(t);\theta)dt = \frac{h}{2}\left\{\mu(X_{ih};\theta) + \mu(X_{(i-1)h};\theta)\right\}.$$

- For the OU process the discrete approximate model is

$$X_{ih} - X_{(i-1)h} = \kappa\mu - \frac{\kappa h}{2}\left(X_{ih} + X_{(i-1)h}\right) + \sigma N(0,h), \quad (41)$$

Parametric methods

- Solving (41) we obtain

$$\begin{aligned} X_{ih} &= \frac{\kappa\mu h}{\left(1+\frac{\kappa h}{2}\right)} + \frac{1-\frac{\kappa h}{2}}{1+\frac{\kappa h}{2}} X_{(i-1)h} + \frac{\sigma}{\left(1+\frac{\kappa h}{2}\right)} N(0,h) \\ &= \kappa\mu h + (1-\kappa h) X_{(i-1)h} + \sigma N(0,h) + O\left(h^{3/2}\right), \end{aligned}$$

- Consistent for parameters in the diffusion function when $h \rightarrow 0$ (in-fill asymptotics) and consistent for parameters in the drift function when $h \rightarrow 0$ and $T \rightarrow \infty$ (both in-fill and long-span asymptotics).

Parametric methods

- Elerian (1998) suggests using the scheme proposed by Milstein (1978). The idea is to take a second order term in a stochastic Taylor series expansion (39):

$$\begin{aligned} X_{ih} &= X_{(i-1)h} + \mu(X_{(i-1)h},\theta)h - g(X_{(i-1)h},\theta)h + \\ &\quad \sigma(X_{(i-1)h},\theta)\sqrt{h}\epsilon_i + g(X_{(i-1)h},\theta)h\epsilon_i^2, \end{aligned} \tag{42}$$

  where

$$g(X_{(i-1)h},\theta) = \frac{1}{2}\sigma'(X_{(i-1)h};\theta)\sigma(X_{(i-1)h};\theta). \tag{43}$$

Parametric methods

- Kessler (1997) advocated approximating the transition density using a Gaussian density with the same mean and variance.
- Nowman (1997) suggested to approximate the model:

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X(t), \theta)dB(t) \tag{44}$$

by

$$dX(t) = \kappa(\mu - X(t))dt + \sigma(X_{(i-1)h}; \theta)dB(t), \quad (i-1)h \leq t < ih. \tag{45}$$

whose exact discrete model is

$$X_{ih} = e^{-\kappa h} X_{(i-1)h} + \mu\left(1 - e^{-\kappa h}\right) + \sigma(X_{(i-1)h}; \theta)\sqrt{\frac{1 - e^{-2\kappa h}}{2\kappa}}\epsilon_i, \tag{46}$$

Finite Sample Theory

- In the context of AR(1) process without intercept, Kendall (1954) obtained the approximation to the bias of ML estimator of ϕ:
$$E(\widehat{\phi}) - \phi \approx -\frac{2\phi}{n} \tag{47}$$
- Since $\kappa = -\ln(\phi)/h$, and the asymptotic variance of $\widehat{\phi}$ is $(1-\phi^2)/n$, we have
$$E(\hat{\kappa}) - \kappa \approx \frac{1}{2T}\left(3 + e^{2\kappa h}\right) \tag{48}$$
- This suggests that the bias is not determined by the sample size (n) but by the time span (T) instead.
- The bias translates into prices.
- The estimation bias can be much more substantial than the bias induced by discretization and mis-specifications; see Phillips and Yu (2005a, 2005b).

Finite Sample Theory

- Let N be the number of observations in the whole sample and decompose the sample into m consecutive subsamples each with ℓ observations, so that $N = m \times \ell$. Phillips and Yu (2005a) proposed the following jackknife estimator of κ, estimator
$$\hat{\kappa}_{jack} = \frac{m}{m-1}\hat{\kappa}_N - \frac{\sum_{i=1}^{m} \hat{\kappa}_{li}}{m^2 - m}, \tag{49}$$
- The bias in $\hat{\kappa}_{jack}$ is not $O(1/T)$ but $O(1/T^2)$. Its variance goes up in finite sample.

Finite Sample Theory

- How to improve jackknife? Simulate-based methods.
- We focus on the OU process. Suppose we need to estimate the parameter κ in the model

$$dX(t) = \kappa(\mu - X(t))dt + \sigma\, dB(t). \tag{50}$$

  from observations $\mathbf{x} = \{X_h, \cdots, X_{Nh}\}$.
- An initial estimator of κ can be obtained, for example, by applying the Euler scheme to $\{X_h, \cdots, X_{Nh}\}$ (call it $\hat{\kappa}_N$). This estimator is subject to both discretization bias and estimation bias.

Parametric methods

- The idea of indirect inference (II) is: given a parameter choice κ, we apply the Euler scheme with a much smaller step size than h (say $\delta = h/10$), which leads to

  $$\tilde{X}^k_{t+\delta} = \kappa(\mu - \tilde{X}^k_t)h + \tilde{X}^k_t + \sigma\sqrt{\delta}\epsilon_{t+\delta}, \tag{51}$$

  where

  $$t = \underbrace{0, \delta, \cdots, h(= 10\delta)}, \underbrace{h + \delta, \cdots, 2h(= 20\delta)}, 2h + \delta, \cdots, Nh. \tag{52}$$

  This sequence may be regarded as a nearly exact simulation from the continuous time OU model for small δ. We then choose every $(h/\delta)^{th}$ observation to form the sequence of $\{\tilde{X}^k_{ih}\}_{i=1}^N$, which can be regarded as data simulated directly from the OU model with the (observationally relevant) step size h.

Parametric methods

- Let $\tilde{\mathbf{x}}^k(\kappa) = \{\tilde{X}_h^k, \cdots, \tilde{X}_{Nh}^k\}$ be data simulated from the true model, where $k = 1, \cdots, K$ with K being the number of simulated paths. It should be emphasized that it is important to choose the number of observations in $\tilde{\mathbf{x}}^k(\kappa)$ to be the same as the number of observations in the observed sequence $\mathbf{x}$ for the purpose of the bias calibration. Another estimator of κ can be obtained by applying the Euler scheme to $\{X_h^k, \cdots, X_{Nh}^k\}$ (call it $\tilde{\kappa}_N^k$). Such an estimator and hence the expected value of them across simulated paths is naturally dependent on the given parameter choice κ.

Parametric methods

- Let $\tilde{\mathbf{x}}^k(\kappa) = \{\tilde{X}_h^k, \cdots, \tilde{X}_{Nh}^k\}$ be data simulated from the true model, where $k = 1, \cdots, K$ with K being the number of simulated paths. It should be emphasized that it is important to choose the number of observations in $\tilde{\mathbf{x}}^k(\kappa)$ to be the same as the number of observations in the observed sequence $\mathbf{x}$ for the purpose of the bias calibration.
- An estimator of κ is obtained, denoted as $\tilde{\kappa}_N^k$.

Parametric methods

- The central idea in II estimation is to match the parameter obtained from the actual data with that obtained from the simulated data. In particular, the II estimator of κ is defined as

$$\hat{\kappa}_{N,K}^{II} = \text{argmin}_{\kappa} \left\| \hat{\kappa}_N - \frac{1}{K} \sum_{h=1}^{K} \tilde{\kappa}_N^k(\kappa) \right\|, \tag{53}$$

- If the sample mean in (53) is replaced with the sample median, the estimator is median unbiased.
- Phillips and Yu (2005a) implemented median unbiased estimation and Phillips and Yu (2009) implemented indirect inference estimation, of κ.
- Both estimators can remove discretization bias and estimation bias. The variance can be smaller than that of MLE in finite sample. In the unit root case, there will be an improvement over MLE even asymptotically.

Finite Sample Theory

- Bootstrap method of Tang and Chen (2009, JoE)

1. Estimate κ using ML and obtain $\hat{\kappa}$.
2. Generate a bootstrap path $\{X_t^*\}_{t=1}^N$ from the estimated model with the sampling interval as the original data.
3. Obtain MLE of κ from the bootstrap sample, $\hat{\kappa}^*$.
4. Repeat Steps 2-3 K times and obtain $\hat{\kappa}^{*,1}, \cdots, \hat{\kappa}^{*,K}$.
5. Calculate $2\hat{\kappa} - \frac{1}{K}\sum_{k=1}^{K} \hat{\kappa}^{*,k}$.

Asymptotic Theory

- There are three alternative ways to develop asymptotic distribution, namely,

$$T = nh, \; n \to \infty, \quad h \text{ fixed} \tag{$A1$}$$

$$T = nh, \; n \to \infty, \quad h \to 0 \text{ and } T \to \infty \tag{$A3$}$$

$$T = nh, \; n \to \infty, \quad T \text{ fixed} \tag{$A2$}$$

- (A1) is the long-span asymptotics. (A2) is in-fill asymptotics. (A3) is the double asymptotics.

Asymptotic Theory

- Examples of (A1): Hermite based ML, Saddlepoint based ML, exact ML, Pedersen (1995).
- Examples of (A2), Euler, Milstein, Nowman, Phillips and Yu (2001), Bergstrom (1966), Bandi and Phillps (2003, 2006).
- Examples of (A3): Bandorff-Nielsen and Shephard (2002).

Asymptotic Theory

- The exact discrete time representation of the Vasicek model with a known mean is

$$X_{th} = \phi X_{(t-1)h} + \sigma \sqrt{\frac{1-e^{-2\kappa h}}{2\kappa}} \epsilon_t, \tag{54}$$

where $\phi = e^{-\kappa h}, \epsilon_t \overset{i.i.d}{\sim} N(0,1)$.

- The least square estimator of ϕ is

$$\hat{\phi}_n = \frac{\sum X_{t-1} X_t}{\sum X_{t-1}^2}$$

If $\kappa > 0$, the model is stationary. Under $(A1)$, by central limit theory of martingale difference sequences, we have $\sqrt{n}(\hat{\phi}_n - \phi) \xrightarrow{d} N(0, 1-\phi^2)$ as $n \to \infty$. By the *Delta* method, under $(A1)$

$$\sqrt{T}(\hat{\kappa}_n - \kappa) \xrightarrow{d} N\left(0, \frac{e^{2\kappa h}-1}{h}\right), \tag{55}$$

Asymptotic Theory

- If $\kappa = 0$, under $(A1)$,

$$T(\hat{\kappa}_n - \kappa) \xrightarrow{d} -\frac{\int_0^1 BdB}{\int_0^1 B^2 dr}. \tag{56}$$

- Define $J_c(r) = \int_0^r e^{c(r-s)} dW(s)$, $\gamma = X_0 / \left(\sigma\sqrt{T}\right)$, $c = -\kappa T$ and

$$A(\gamma, c) = \gamma \int_0^1 e^{cr} dW(r) + \int_0^1 J_c(r) dW(r)$$

$$B(\gamma, c) = \gamma^2 (e^{2c} - 1)/2c + 2\gamma \int_0^1 e^{cr} J_c(r) dW(r) + \int_0^1 J_c^2(r) dW(r).$$

- Regardless of κ, under $(A2)$,

$$T(\hat{\kappa}_n - \kappa) \xrightarrow{d} -\frac{A(\gamma, c)}{B(\gamma, c)}. \tag{57}$$

Asymptotic Theory

- Under $(A3)$ with $T \to \infty$ and $h \to 0$, the asymptotic distribution is

  $$\sqrt{T}(\hat{\kappa}_n - \kappa) \xrightarrow{d} N(0, 2\kappa) \tag{58}$$

  for $\kappa > 0$ and

  $$T(\hat{\kappa}_n - \kappa) \xrightarrow{d} -\frac{\int_0^1 W dW}{\int_0^1 W^2 dr}. \tag{59}$$

  for $\kappa = 0$.

Finite Sample Theory

- The simple bias approximation of κ in the Vasicek model with a known mean and the initial condition $X_0 \sim N(0, \sigma^2/(2\kappa))$ is

$$E(\hat{\kappa}) - \kappa \approx \frac{1}{2T}\left(3 + e^{2\kappa h}\right)$$

- Yu (2009) showed that this formula works terribly in the near unit root case. But, in practice, κ is very close to 0.
- Yu (2009) obtained a better approximation to the bias of ML estimator of κ:

$$E(\hat{\kappa}) - \kappa \approx \frac{1}{2T}\left(3 + e^{2\kappa h}\right) - \frac{2(1 - e^{-2n\kappa h})}{Tn(1 - e^{-2\kappa h})}. \tag{60}$$

- The correction term is not negligible when κ or h is close to 0.

Finite Sample Theory

- In the context of OU process with an unknown mean, Tang and Chen (2009) obtained an approximation to the bias of ML estimator of κ:

$$E(\hat{\kappa}) - \kappa \approx \frac{1}{2T}(5 + e^{2\kappa h} + 2e^{\kappa h})$$

- This is analogue to the bias formula $\frac{1+3\phi}{n}$ for the AR(1) model with the intercept derived by Kendall (1954).
- Tang and Chen (2009) also obtained the bias of Nowman's ML estimator of κ in the CIR model:

$$E(\hat{\kappa}) - \kappa \approx \frac{1}{2T}(5 + e^{2\kappa h} + 2e^{\kappa h})$$