Parametric Estimation of SDE

伊藤克哉

東京大学

2016/10/31

講義ノート スライド コードは
https://github.com/KatsuyaITO/NSofSDE

Lamperti transform

以下では

$$dX_t = b(t, X_t)dt + \sigma(X_t)dW_t$$

という SDE の解 X_t について考える. このとき,

$$Y_t = F(X_t) = \int_z^{X_t} \frac{1}{\sigma(u)} du$$

という変換を施すと,

$$dY_t = b_Y(t, Y_t)dt + dW_t$$

とできる.

Shoji-Ozaki method

Shoji-Ozaki 法

$$dX_t = b(t, X_t)dt + \sigma dW_t$$

という SDE を $[s, s + \Delta s)$ 上で局所的に近似

$$dX_t = (L_s X_t + tM_s + N_s)dt + \sigma dW_t$$

ただし,

$$L_s = b_x(s, X_s),\ M_s = \frac{\sigma^2}{2} b_{xx}(s, X_s) + b_t(s, X_s)$$

$$N_s = b(s, X_s) - X_s b_x(s, X_s) - sM_s$$

Shoji-Ozaki method

X_t に対して,$Y_t = e^{-L_s t} X_t$ 変換することによって,

$$Y_t = Y_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dW_u$$

という SDE が得られる.

Shoji-Ozaki method

X_t に対して,$Y_t = e^{-L_s t}X_t$ 変換することによって,

$$Y_t = Y_s + \int_s^t (M_s u + N_s)e^{-L_s u}du + \sigma \int_s^t e^{-L_s u}dW_u$$

という SDE が得られる.
これを部分的に解いて,X_t に戻す事によって,

$$X_{s+\Delta s} = A(X_s)X_s + B(X_s)Z$$

ただし,

$$A(X_s) = 1 + \frac{b(s,X_s)}{X_s L_s}(e^{L_s \Delta s} - 1) + \frac{M_s}{X_s L_s^2}(e^{L_s \Delta s} - 1 - L_s \Delta s) \quad (1.1)$$

$$B(X_s) = \sigma \sqrt{\frac{\exp 2L_s \Delta s - 1}{2L_s}}, \; Z \sim N(0,1) \quad (1.2)$$

Exact Sampling(基本的な場合の棄却法)

g という確率密度関数をもつ確率変数の生成法を知っている.
f という確率密度関数をもつ確率変数を生成したい

$$f(x) \leq \epsilon g(x)$$

というような $\epsilon > 0$ が存在しているとする.
このとき次のアルゴリズムに従えば良い.

1. g に従う数 y を 1 つ生成する.
2. 一様分布 $U(0,1)$ に従う乱数 u を 1 つ生成する.
3. $u < \epsilon f(y)/g(y)$ を満たすならば,y を採用し, それ以外ならば 1 に戻る.

Exact Sampling(確率微分方程式に応用)

b が x で偏微分可能であり,$\frac{1}{2}b^2(x) + \frac{1}{2}b_x(x)$ が上にも下にも有界であると仮定する.
ここで,$k_1 \le \frac{1}{2}b^2(x) + \frac{1}{2}b_x(x) \le k_2$ であるとする.

$$\phi(x) = \frac{1}{2}b^2(x) + \frac{1}{2}b_x(x) - k_1, \ M = k_2 - k_1$$

とおく.

$$h(z) = \exp\left\{ \int_0^z b(u)du - \frac{(z-x)^2}{2\Delta} \right\}$$

を規格化して,$\hat{h}$ とおく. この時次のアルゴリズムに従えば良い.

Exact Sampling

1. $\hat{h}$ に従うような y を生成する.
2. 0 のとき 0 で,Δ のとき y のブラウン橋を作る.
3. k を $\lambda = \Delta M$ の Poisson 過程から作る.
4. $[0, T] \times [0, M]$ 上の一様分布の実現 $\{x_1, \cdots, x_k\}$,$x_i = (t_i, v_i)$ を得る.
5. 全ての i について $\phi(y_i) \leq v_i$ ならば採用し, それ以外ならば 1 に戻る.

Continuous な場合

$$dX_t = b_\theta(X_t)dt + \sigma_\theta(X_t)dW_t$$

θ によりパラメトライズされた自励系の SDE.
$\theta \in \Theta \subset \mathbb{R}^p$ は p 個のパラメータであり,θ_0 を推測する.

ergodic な場合

過程がエルゴード性を満たし, 連続的である場合には容易に最尤法を適用することができる.
連続的な場合は実務上は起こりにくいが, 理論的には重要である

ergodic とは

定義 2.1

X_t という過程が *ergodic* であるとは,$\exists\xi$:確率変数,$\forall h(x)$:可測関数

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T h(X_t)dt \to E[h(\xi)]$$

が成立すること.
このとき,ξ の分布関数を π と置いて,**invariant density** または *stationary density* という.

ergodic の特徴付け

定理 2.2

recurrent で *positive* な過程は *ergodic* であり，このとき *invariant density* π は

$$\pi(x) = \frac{1}{G\sigma(x)^2} \exp\left\{ 2 \int_0^x \frac{b(y)}{\sigma(y)^2} dy \right\}$$

として表される．

recurrent と positive とは

定義 2.3

$$\tau_a = \inf\{t \geq 0 | X_t = a\}, \tau_{ab} = \inf\{t \geq \tau_a | X_t = b\}$$

と定める. 過程 X_t が **recurrent** であるとは,$P(\tau_{ab} < \infty) = 1$ が任意の $a, b \in \mathbb{R}$ に対して成り立つことである.
recurrent な過程 X_t が **positive** であるということは,$E[\tau_{ab}] < \infty$ が任意の $a, b \in \mathbb{R}$ に対して成り立つことである.

recurrent と positive に対する特徴付け

命題 2.4

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

という過程が *recurrent* である $\Leftrightarrow$

$$V(x) = \int_0^x \exp\left(-2\int_0^y \frac{b(u)}{\sigma(u)^2}du\right)dy \to \pm\infty \quad x \to \pm\infty$$

recurrent な過程が *positive* である $\Leftrightarrow$

$$G = \int_{-\infty}^{\infty} \sigma(y)^{-2} \exp\left(2\int_0^y \frac{b(u)}{\sigma(u)^2}du\right)dy < \infty$$

ergodic な場合の最尤推定

任意の θ に対して ergodic な過程を考える.

ergodicな場合の最尤推定

任意の θ に対して ergodic な過程を考える. まず quadratic variation を計算して,

$$< X, X >_t = \lim_{n\to\infty} \sum_{k=1}^{2^n} (X_{t\wedge(k/2^n)} - X_{t\wedge((k-1)/2^n)})^2 = \int_0^t \sigma^2(X_s, \theta) ds$$

σ のパラメータを推定する.

ergodicな場合の最尤推定

任意のθに対してergodicな過程を考える. まずquadratic variationを計算して,

$$< X, X >_t = \lim_{n\to\infty} \sum_{k=1}^{2^n} (X_{t\wedge(k/2^n)} - X_{t\wedge((k-1)/2^n)})^2 = \int_0^t \sigma^2(X_s, \theta) ds$$

σのパラメータを推定する.
その後,Girsanovの定理によって, 次のように尤度が表される.

$$L_T(\theta) = \frac{dP_1}{dP}(X) = \exp\left(\int_0^T \frac{b_\theta(X_s)}{\sigma^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{b_\theta^2(X_s)}{\sigma^2(X_s)} ds \right)$$

つまりこの尤度を最大にするようなθを求めれば良い.

Girsanov の定理

定理 2.5 (Girsanov)

$$dX_t = b_1(X_t)dt + \sigma(X_t)dW_t,\ X_0 = X_0^{(1)} \tag{2.1}$$
$$dX_t = b_2(X_t)dt + \sigma(X_t)dW_t,\ X_0 = X_0^{(2)} \tag{2.2}$$
$$dX_t = \sigma(X_t)dW_t,\ X_0 = X_0 \tag{2.3}$$

という $t \in [0, T]$ 上で定義された *3* つの伊藤過程を考える*.*
この過程から引き起こされる確率測度を P_1, P_2, P_3 とおく*.*
ここで*,* 各 b と σ は *Global Lipschitz* であり*,Linear growth* であると仮定する *(*ゆえに各確率微分方程式の解が存在する*)*
さらに*,*f_1, f_2, f という初期値たちの密度関数は同じサポートを持つと仮定する*.*

Girsanov の定理

このとき次のようにしてこれらの Radon-Nikodym 微分が存在して, 次のように表される.

$$\frac{dP_1}{dP}(X) = \frac{f_1(X_0)}{f(X_0)} \exp\left\{ \int_0^T \frac{b_1(X_s)}{\sigma^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{b_1^2(X_s)}{\sigma^2(X_s)} ds \right\}$$

$$\frac{dP_2}{dP_1}(X) = \frac{f_2(X_0)}{f_1(X_0)} \exp\left\{ \int_0^T \frac{b_2(X_s) - b_1(X_s)}{\sigma^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{b_2^2(X_s) - b_1^2(X_s)}{\sigma^2(X_s)} ds$$

Exact likelihood inference

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t,\ X_0 = x_0,\ \theta_1, \theta_2 \in \mathbb{R}, \theta_3 > 0$$

であるような Ornstein-Uhlenbeck 過程に対して, 条件付き密度関数は

$$p_\theta(t, \cdot|x) \sim N(m(t,x), v(t,x))$$

$$m(t,x) = \frac{\theta_1}{\theta_2} + (x_0 - \frac{\theta_1}{\theta_2})e^{-\theta_2 t}$$

$$v(t,x) = \frac{\theta_3^2(1 - e^{-2\theta_2 t})}{2\theta_2}$$

という平均と分散をもつ正規分布に従う. 故にこれにより直接尤度を計算することができる.

geometric Brownian motion

次のように表される geometric Brownian motion を考える.

$$dX_t = \theta_1 X_t dt + \theta_2 X_t dW_t,\ X_0 = x_0,\ \theta_1 \in \mathbb{R},\ \theta_2 > 0$$

同様に, このモデルの条件付き密度関数は,

$$p_\theta(t, \cdot | x) \sim LN(m(t, x), v(t, x))$$

$$m(t, x) = xe^{\theta_1 t},\ v(t, x) = x^2 e^{2\theta_1 t}(e^{\theta_2^2 t} - 1)$$

というような対数正規分布によって表される. 故にこれにより直接尤度を計算することができる.

Cox-Ingersoll-Ross Model

次のように表される Cox-Ingersoll-Ross モデルを考える.

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} dW_t,\ X_0 = x_0 > 0,\ \theta_1, \theta_2 \theta_3 > 0$$

同様に, このモデルの条件付き密度関数は,

$$p_\theta(t, \cdot|x) = c \exp -u - v \left(\frac{u}{v}\right)^{q/2} I_q(2\sqrt{uv}),\ x, y > 0$$

という風に表される. ただし,

$$c = \frac{2\theta_2}{\theta_2^3(1 - e^{-\theta_2 t})}, q = \frac{2\theta_1}{\theta_3^2} - 1$$

$$u = cx \exp -\theta_2 t, v = cy$$

であり, I_q は修正 Bessel 関数である.

$$I_q(x) = \sum_{k=0}^{\infty} \left(\frac{x}{2}\right)^{2k+q} \frac{1}{k!\Gamma(k+q+1)}$$

Pseudo-likelihood methods

SDE の離散近似スキームを使ってパスを生成し, 過程の尤度を求める方法を **Pseudo-likelihood 法**という.
ここでは,

1. Euler method
2. Elerian method
3. Shoji-Ozaki method

について考える.

Euler method

$$dX_t = b_\theta(X_t)dt + \sigma_\theta(X_t)dW_t$$

という確率微分方程式に対して Euler-Maruyama 近似を適用すると,

$$\Delta X = b_\theta(X_t)\Delta t + \sigma_\theta(X_t)\Delta W_t$$

であるので,ΔX はそれぞれ, 平均 $b_\theta(X_t)$ 分散 $\sigma_\theta^2(X_t)$ の正規分布に従っている.

Euler method

故に,Euler-Maruyama 近似が強収束するとき, この対数尤度 $l_n(\theta)$ は

$$l_n(\theta) = -\frac{1}{2}\left\{\sum_{i=1}^{n} \frac{X_i - X_{i-1} - b(X_{i-1}, \theta)\Delta)^2}{\sigma^2 \Delta} + n \log(2\pi\sigma^2\Delta)\right\}$$

という式によって表される.
N.Yoshida(1992) によると,

$$\hat{\sigma}^2 = \frac{1}{n\Delta}\sum_{i=1}^{n}(X_i - X_{i-1})^2$$

によって σ を推定することができる.

Exact Likelihood v.s. Euler

ここで, 最初に述べた Exact likelihood と Euler approximation の性質を見てみる.

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t,\ X_0 = x_0,\ \theta_1, \theta_2 \in \mathbb{R}, \theta_3 > 0$$

という Ornstein-Uhlenbeck 過程に対して 2 つの方法を適用しすると,

Exact Likelihood v.s. Euler

ここで, 最初に述べた Exact likelihood と Euler approximation の性質を見てみる.

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t,\ X_0 = x_0,\ \theta_1, \theta_2 \in \mathbb{R}, \theta_3 > 0$$

という Ornstein-Uhlenbeck 過程に対して 2 つの方法を適用しすると,

$$m(\Delta, x) = x\exp(-\theta_2\Delta) + \frac{\theta_1}{\theta_2}(1 - \exp(-\theta_2\Delta)),\ v(\Delta, x) = \frac{\theta_3^2(1 - \exp^{-2\theta_2\Delta})}{2\theta_2}$$

という平均と分散が得られ, また Euler approximation を行うと,

Exact Likelihood v.s. Euler

ここで, 最初に述べた Exact likelihood と Euler approximation の性質を見てみる.

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t,\ X_0 = x_0,\ \theta_1, \theta_2 \in \mathbb{R}, \theta_3 > 0$$

という Ornstein-Uhlenbeck 過程に対して 2 つの方法を適用すると,

$$m(\Delta, x) = x\exp(-\theta_2\Delta) + \frac{\theta_1}{\theta_2}(1 - \exp(-\theta_2\Delta)),\ v(\Delta, x) = \frac{\theta_3^2(1 - \exp^{-2\theta_2\Delta})}{2\theta_2}$$

という平均と分散が得られ, また Euler approximation を行うと,

$$m(\Delta, x) = x(1 - \theta_2\Delta) + \theta_1\Delta,\ v(\Delta, x) = \theta_3^2\Delta$$

という平均と分散が得られる. これらは,$\Delta \to 0$ では一致する.

Exact Likelihood v.s. Euler

一方で,

$$dX_t = \theta_1 X_t dt + \theta_2 X_t dW_t,\ X_0 = x_0,\ \theta_1 \in \mathbb{R},\ \theta_2 > 0$$

という geometric Brownian motion を考える. この時, 上の式出てきた最尤法を当てはめると,

Exact Likelihood v.s. Euler

一方で,

$$dX_t = \theta_1 X_t dt + \theta_2 X_t dW_t,\ X_0 = x_0,\ \theta_1 \in \mathbb{R},\ \theta_2 > 0$$

という geometric Brownian motion を考える. この時, 上の式出てきた最尤法を当てはめると,

$$\hat{\theta}_{1,n} = \frac{1}{n\Delta}\sum_{i=1}^{n}\left(\frac{X_i}{X_{i-1}} - 1\right), \hat{\theta}_{2,n} = \frac{1}{n\Delta}\sum_{i=1}^{n}\left(\frac{X_i}{X_{i-1}} - 1 - \hat{\theta}_{1,n}^2\Delta\right),$$

という最尤推定値を得る. ここで,Δ が無視できないほど大きい場合を考えると,

Exact Likelihood v.s. Euler

一方で,

$$dX_t = \theta_1 X_t dt + \theta_2 X_t dW_t,\ X_0 = x_0,\ \theta_1 \in \mathbb{R},\ \theta_2 > 0$$

という geometric Brownian motion を考える. この時, 上の式出てきた最尤法を当てはめると,

$$\hat{\theta}_{1,n} = \frac{1}{n\Delta}\sum_{i=1}^{n}\left(\frac{X_i}{X_{i-1}} - 1\right), \hat{\theta}_{2,n} = \frac{1}{n\Delta}\sum_{i=1}^{n}\left(\frac{X_i}{X_{i-1}} - 1 - \hat{\theta}_{1,n}^2\Delta\right),$$

という最尤推定値を得る. ここで,Δ が無視できないほど大きい場合を考えると,

$$\hat{\theta}_{1,n} \to \frac{1}{\Delta}(e^{\theta_1\Delta} - 1) \neq \theta_1, \hat{\theta}_{2,n} \to \frac{1}{\Delta}e^{2\theta_1\Delta}(e^{\theta_2^2\Delta} - 1) \neq \theta_2$$

という確率収束が得られるので,Δ の影響は大きい.

Elerian method

Milstein 近似により,

$$\Delta X_t = b(t, X_t)\Delta t + \sigma(t, X_t)\Delta W + \frac{1}{2}\sigma(t, X_t)\sigma_x(t, X_t)((\Delta W)^2 - \Delta t)$$

定理 2.6

Milstein 近似がオーダー 1 で強収束しているとき

$$p(t, y|x) = \frac{z^{-1/2}\cosh(\sqrt{Cz})}{|A|\sqrt{2\pi}}\exp(-\frac{C+z}{2})$$

として得られる. ただし,

$$A = \frac{\sigma(x)\sigma_x(x)t}{2}, B = -\frac{\sigma(x)}{2\sigma_x(x)} + x + b(x)t - A,$$

$$z = \frac{y-B}{A}, C = \frac{1}{\sigma_x^2(x)t}$$

Shoji-Ozaki method

Shoji-Ozaki 法

$$dX_t = b(t, X_t)dt + \sigma dW_t$$

という SDE を $[s, s + \Delta s)$ 上で局所的に近似

$$dX_t = (L_s X_t + tM_s + N_s)dt + \sigma dW_t$$

ただし,

$$L_s = b_x(s, X_s),\ \ M_s = \frac{\sigma^2}{2} b_{xx}(s, X_s) + b_t(s, X_s)$$

$$N_s = b(s, X_s) - X_s b_x(s, X_s) - sM_s$$

Shoji-Ozaki method

X_t に対して,$Y_t = e^{-L_s t} X_t$ 変換することによって,

$$Y_t = Y_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dW_u$$

という SDE が得られる.

Shoji-Ozaki method

X_t に対して, $Y_t = e^{-L_s t} X_t$ 変換することによって,

$$Y_t = Y_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dW_u$$

という SDE が得られる.
これを部分的に解いて, X_t に戻す事によって,

$$X_{s+\Delta s} = A(X_s) X_s + B(X_s) Z$$

ただし,

$$A(X_s) = 1 + \frac{b(s, X_s)}{X_s L_s}(e^{L_s \Delta s} - 1) + \frac{M_s}{X_s L_s^2}(e^{L_s \Delta s} - 1 - L_s \Delta s) \quad (2.4)$$

$$B(X_s) = \sigma \sqrt{\frac{\exp 2L_s \Delta s - 1}{2L_s}}, \; Z \sim N(0, 1) \quad (2.5)$$

である. これを利用することによって, 尤度関数を計算することができる.

Approximated likelihood methods

次に, パスを近似するのではなく直接尤度を近似する **Approximated likelihood methods** について考える.

1. Kessler method
2. Simulated likelihood method
3. Shoji-Ozaki method

について考える.

Kessler method

$$dX_t = b_\theta(X_t)dt + \sigma_\theta(X_t)dW_t$$

というような自励系の ergodic な確率過程を考える.
ここで

$$t_i^n = i\Delta_n$$

とおいて, この幅で離散化することを考える.
ここで,Ito-Taylor 展開を考え,

$$r_l(\Delta_n, x, \alpha) = \sum_{i=0}^{l} \frac{\Delta_n^i}{i!} L_\alpha^i f(x)$$

という第 l 項まで展開したものを考える.

Kessler method

このとき,$n\Delta_n^p \to 0$ であるとする. $k_0 = [p/2]$ と定めると次のようにして,尤度を定めることができる.

$$l_{p,n}(\alpha) = \sum_{i=1}^{n} \frac{(X_{t_i^n} - r_{k_0}(\Delta_n, X_{t_{i-1}^n}, \alpha))^2}{\Delta_n \sigma_{i-1}^2(\theta)} \left\{ 1 + \sum_{j=1}^{k_0} \Delta_n^j d_j(X_{t_{i-1}^n}, \alpha) \right\}$$

$$+ \sum_{i=1}^{n} \left\{ \log \sigma_{i-1}^2(\theta) + \sum_{j=1}^{k_0} + \Delta_n^j e_j(X_{t_{i-1}^n}, \alpha) \right\}$$

ただし,d_j, e_j は特別な関数の Taylor 展開をしたときの係数である.

Simulated likelihood method

$p_\theta(\Delta, y|x)$ を求めることを考える.$\delta \ll \Delta$ とする. このとき,

$$p_\theta(\Delta, y|x) = \int p_\theta(\delta, y|z) p_\theta(\Delta - \delta, z|x) dz = E_z[p_\theta(\delta, y|z)|\Delta - \delta]$$

であることがわかる. ここで, 右の期待値を Monte Carlo 法によって求めると, $p_\theta(\Delta - \delta, z|x)$ は正規分布であるので,

$$\hat{p}_\theta^{(N,M)}(\Delta, X_{t+\Delta}|x) = \frac{1}{M} \sum_{i=1}^{M} \phi_\theta(\delta, X_{t+\Delta}|z_i)$$

として, 確率密度関数を求めることができる.

Hermite polynomials expansion of the likelihood

$$H_j(z) = e^{\frac{z^2}{2}} \frac{d^j}{dz^j} e^{-\frac{z^2}{2}}$$

という修正 Hermite 多項式を考える.

$$F(\gamma) = \int_0^{\gamma} \frac{du}{\sigma_\theta(u)}$$

という Lamperti 変換を加えることによって, まず

$$dY_t = \mu_Y(Y_t, \theta)dt + dW_t$$

という風に変換する. その次に,

$$Z = \Delta^{-1/2}(Y - y_0)$$

という変換を施す.

Hermite polynomials expansion of the likelihood

$$p_Z(\Delta, z|y_0, \theta) = \Delta^{1/2} p_Y(\Delta, \Delta^1 2z + y_0|y_0, \theta)$$

という確率密度関数を求めることを考える. ここで,$H_j(z)/\sqrt{j!}$ は L^2 空間の直交基底であることを使うと,p_Z を分解することができる. そして, その展開を途中で打ち切ることによって, 尤度の近似を行う.