Efficient estimation for ergodic diffusions sampled at high frequency

Michael Sørensen

Department of Mathematical Sciences
University of Copenhagen, Denmark

http://www.math.ku.dk/∼michael

Discretely observed diffusion

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad \theta \in \Theta \subseteq \mathbb{R}^p$$

Data: $X_{t_1}, \cdots, X_{t_n}$, $t_1 < \cdots < t_n$.

Discretely observed diffusion

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad \theta \in \Theta \subseteq \mathbb{R}^p$$

Data: $X_{t_1}, \cdots, X_{t_n}, \quad t_1 < \cdots < t_n$.

Review papers:

Helle Sørensen (2004)   Int. Stat. Rev.

Bibby, Jacobsen and Sørensen (2004)

Sørensen (2007)

Likelihood inference

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad \theta \in \Theta \subseteq \mathbb{R}^p$$

Data: $X_{t_1}, \cdots, X_{t_n}$, $t_1 < \cdots < t_n$.

Likelihood-function:

$$L_n(\theta) = \prod_{i=1}^{n} p(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta),$$

where $t_0 = 0$ and $\Delta_i = t_i - t_{i-1}$.

$y \mapsto p(\Delta, x, y; \theta)$ is the probability density function of the conditional distribution of $X_{t+\Delta}$ given that $X_t = x$.

Likelihood inference

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad \theta \in \Theta \subseteq \mathbb{R}^p$$

Data: $X_{t_1}, \cdots, X_{t_n}, \quad t_1 < \cdots < t_n$.

Likelihood-function:

$$L_n(\theta) = \prod_{i=1}^{n} p(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta),$$

where $t_0 = 0$ and $\Delta_i = t_i - t_{i-1}$.

$y \mapsto p(\Delta, x, y; \theta)$ is the probability density function of the conditional distribution of $X_{t+\Delta}$ given that $X_t = x$.

Score function:

$$U_n(\theta) = \partial_\theta \log L_n(\theta) = \sum_{i=1}^{n} \partial_\theta \log p(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta).$$

Under weak regularity conditions, the score function is a P_θ-martingale

Quadratic estimating functions

Approximate likelihood function

$$L_n(\theta) \doteq M_n(\theta) = \prod_{i=1}^{n} q(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta)$$

$$p(\Delta, x, y; \theta) \doteq q(\Delta, x, y; \theta) = \frac{1}{\sqrt{2\pi\Phi(\Delta, x; \theta)}} \exp\left[\frac{(y - F(\Delta, x; \theta))^2}{2\Phi(\Delta, x; \theta)}\right]$$

$$F(x;\theta) = E_\theta(X_\Delta | X_0 = x) \quad \text{and} \quad \Phi(x;\theta) = \text{Var}_\theta(X_\Delta | X_0 = x)$$

Quadratic estimating functions

Approximate likelihood function

$$L_n(\theta) \doteq M_n(\theta) = \prod_{i=1}^{n} q(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta)$$

$$p(\Delta, x, y; \theta) \doteq q(\Delta, x, y; \theta) = \frac{1}{\sqrt{2\pi\Phi(\Delta, x; \theta)}} \exp\left[\frac{(y - F(\Delta, x; \theta))^2}{2\Phi(\Delta, x; \theta)}\right]$$

$$F(x; \theta) = E_\theta(X_\Delta | X_0 = x) \quad \text{and} \quad \Phi(x; \theta) = \text{Var}_\theta(X_\Delta | X_0 = x)$$

Approximate score function

$$\partial_\theta \log M_n(\theta) = \sum_{i=1}^{n} \left\{ \frac{\partial_\theta F(\Delta_i, X_{t_{i-1}}; \theta)}{\Phi(\Delta_i, X_{t_{i-1}}; \theta)} [X_{t_i} - F(\Delta_i, X_{t_{i-1}}; \theta)] \right.$$

$$\left. + \frac{\partial_\theta \Phi(\Delta_i, X_{t_{i-1}}; \theta)}{2\Phi(\Delta_i, X_{t_{i-1}}; \theta)^2} [(X_{t_i} - F(\Delta_i, X_{t_{i-1}}; \theta))^2 - \Phi(\Delta_i, X_{t_{i-1}}; \theta)] \right\}$$

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) \quad = \quad \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

$\uparrow$ p-dimensional $\uparrow$ real valued

Transition operator: $\pi_\theta^\Delta f(x; \theta) = E_\theta(f(X_\Delta; \theta) \,|\, X_0 = x)$

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) \quad = \quad \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

↑ p-dimensional ↑ real valued

$G_n(\theta)$ is a P_θ-martingale:

$$E_\theta(a_j(X_{t_{i-1}}, \Delta_i; \theta)[f_j(X_{t_i}; \theta) - \pi_\theta^{\Delta_i} f_j(X_{t_{i-1}}; \theta)] \mid X_{t_1}, \cdots, X_{t_{i-1}}) = 0$$

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) \quad = \quad \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

$$\uparrow \qquad\qquad\qquad \uparrow$$

p-dimensional   real valued

$G_n(\theta)$ is a P_θ-martingale:

$$E_\theta(a_j(X_{t_{i-1}}, \Delta_i; \theta)[f_j(X_{t_i}; \theta) - \pi_\theta^{\Delta_i} f_j(X_{t_{i-1}}; \theta)] \mid X_{t_1}, \cdots, X_{t_{i-1}}) = 0$$

G_n–estimator(s): $G_n(\hat{\theta}_n) = 0$

Bibby and Sørensen (1995,1996)

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

- Easy asymptotics by martingale limit theory

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

- Easy asymptotics by martingale limit theory

- Simple expression for Godambe-Heyde optimal estimating function

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

- Easy asymptotics by martingale limit theory
- Simple expression for Godambe-Heyde optimal estimating function
- Approximates the score function, which is a P_θ-martingale

Martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

- Easy asymptotics by martingale limit theory
- Simple expression for Godambe-Heyde optimal estimating function
- Approximates the score function, which is a P_θ-martingale
- Particular and most efficient instance of GMM

Asymptotics - ergodic diffusions

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad v(x;\theta) = \sigma^2(x;\theta)$$

Asymptotics - ergodic diffusions

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad v(x;\theta) = \sigma^2(x;\theta)$$

Assume that

$$\int_{x^\#}^{r} s(x;\theta)dx = \int_{\ell}^{x^\#} s(x;\theta)dx = \infty \qquad \text{and} \qquad \int_{\ell}^{r} \tilde{\mu}_\theta(x)dx = A(\theta) < \infty,$$

where $x^\#$ is an arbitrary point in (ℓ, r)

$$s(x;\theta) = \exp\left(-2\int_{x^\#}^{x} \frac{b(y;\theta)}{v(y;\theta)}dy\right) \qquad \text{and} \qquad \tilde{\mu}_\theta(x) = [s(x;\theta)v(x;\theta)]^{-1}$$

Asymptotics - ergodic diffusions

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t \qquad v(x;\theta) = \sigma^2(x;\theta)$$

Assume that

$$\int_{x^\#}^{r} s(x;\theta)dx = \int_{\ell}^{x^\#} s(x;\theta)dx = \infty \qquad \text{and} \qquad \int_{\ell}^{r} \tilde{\mu}_\theta(x)dx = A(\theta) < \infty,$$

where $x^\#$ is an arbitrary point in (ℓ, r)

$$s(x;\theta) = \exp\left(-2\int_{x^\#}^{x} \frac{b(y;\theta)}{v(y;\theta)}dy\right) \qquad \text{and} \qquad \tilde{\mu}_\theta(x) = [s(x;\theta)v(x;\theta)]^{-1}$$

X is ergodic with invariant measure $\mu_\theta(x) = \tilde{\mu}_\theta(x)/A(\theta)$

$$Q_\theta^\Delta(x,y) = \mu_\theta(x)p(\Delta, x, y;\theta)$$

Asymptotics - low frequency

Assume that $t_i = \Delta i$ and the identifiability condition that

$$Q_{\theta_0}^{\Delta}\left(g(\Delta, \theta)\right) = 0 \quad \text{if and only if} \quad \theta = \theta_0$$

and weak regularity conditions.

Asymptotics - low frequency

Assume that $t_i = \Delta i$ and the identifiability condition that

$$Q_{\theta_0}^{\Delta}\left(g(\Delta,\theta)\right) = 0 \quad \text{if and only if} \quad \theta = \theta_0$$

and weak regularity conditions.

Then a consistent estimator $\hat{\theta}_n$ that solves the estimating equation $G_n(\theta) = 0$ exists and is unique in any compact subset of Θ containing θ_0 with a probability that goes to one as $n \to \infty$. Moreover,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N\left(0, S_{\theta_0}^{-1} V_{\theta_0} (S_{\theta_0}^T)^{-1}\right)$$

under P_{θ_0}, where

$$V_\theta = Q_{\theta_0}^{\Delta}\left(g(\Delta,\theta) g(\Delta,\theta)^T\right) \quad \text{and} \quad S_\theta = \left\{Q_{\theta_0}^{\Delta}\left(\partial_{\theta_j} g_i(\Delta;\theta)\right)\right\}$$

GMM estimator

$\tilde{\theta}_n$ is found by minimizing

$$K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)\right]$$

where

$$W_n = \frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \bar{\theta}_n) g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \bar{\theta}_n)^T$$

and $\bar{\theta}_n$ is a consistent estimator.

GMM estimator

$\tilde{\theta}_n$ is found by minimizing

$$K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)\right]$$

where

$$W_n = \frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \bar{\theta}_n) g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \bar{\theta}_n)^T$$

and $\bar{\theta}_n$ is a consistent estimator.

$$\partial_\theta K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} \partial_\theta g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)\right] = 0$$

GMM estimator

$$\partial_\theta K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} \partial_\theta g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)\right]$$

$$\frac{1}{n}\sum_{i=1}^{n} \partial_{\theta^T} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta) \to S_\theta \qquad\qquad W_n \to V_\theta$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)$$

GMM estimator

$$\partial_\theta K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n}\partial_\theta g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)\right]$$

$$\frac{1}{n}\sum_{i=1}^{n}\partial_{\theta^T} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta) \to S_\theta \qquad W_n \to V_\theta$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} A_\theta(X_{t_{i-1}})(f_\theta(X_{t_i}) - \pi_\theta^\Delta f_\theta(X_{t_{i-1}}))$$

GMM estimator

$$\partial_\theta K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} \partial_\theta g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)\right]$$

$$\frac{1}{n}\sum_{i=1}^{n} \partial_{\theta^T} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta) \to S_\theta \qquad W_n \to V_\theta$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} A_\theta(X_{t_{i-1}}) (f_\theta(X_{t_i}) - \pi_\theta^\Delta f_\theta(X_{t_{i-1}}))$$

GMM estimator

$$\partial_\theta K_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} \partial_\theta g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)^T\right] W_n^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)\right]$$

$$\frac{1}{n}\sum_{i=1}^{n} \partial_{\theta^T} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta) \to S_\theta \qquad\qquad W_n \to V_\theta$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} g(\Delta_i, X_{t_i}, X_{t_{i-1}};\theta)$$

$$\partial_\theta K_n(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} S_\theta^T V_\theta^{-1} A_\theta(X_{t_{i-1}})(f_\theta(X_{t_i}) - \pi_\theta^\Delta f_\theta(X_{t_{i-1}}))$$

Hansen (1982, 1985)

Approximate martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

$$\pi_\theta^\Delta f_j(x; \theta) = E_\theta(f_j(X_\Delta; \theta) \mid X_0 = x)$$

Approximate martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)]$$

$$\pi_\theta^\Delta f_j(x; \theta) = E_\theta(f_j(X_\Delta; \theta) \,|\, X_0 = x)$$

$$\pi_\theta^\Delta f_j(x; \theta) = f_j(x; \theta) + \Delta \left\{ b(x; \theta)\partial_x f_j(x; \theta) + \tfrac{1}{2}\sigma^2(x; \theta)\partial_x^2 f_j(x; \theta) \right\} + O(\Delta^2)$$

Yoshida (1992)

Kessler (1997)

Kelly, Platen and Sørensen (2004)

Jacobi diffusion

Larsen & Sørensen (2007):

$$dX_t = -\beta[X_t - (m + \gamma z)]dt + \sigma\sqrt{z^2 - (X_t - m)^2}dW_t$$

The eigenfunctions are given in terms of Jacobi polynomials

Asymptotic information at $(\beta, \gamma, \sigma^2) = (0.02, 0, 0.01)$:

Eigenfunction no.	1	2	1 & 2
Inf. for $\hat{\beta}$	47.4	44.8	49.2
Inf. for $\hat{\sigma}^2$	0	759	5016

For optimal estimating functions based on more than two eigenfunctions, the information is not increased by more than 1 - 3 per cent

High frequency asymptotics

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

$\theta = (\alpha,\beta) \in \Theta \subseteq \mathbb{R}^2$ θ_0 is the true parameter value

State space: (ℓ, r)

Ergodic with invariant measure μ_θ.

Data: $X_{t_0^n}, \ldots, X_{t_n^n}$ $t_i^n = i\Delta_n$, $i = 0, \ldots, n$.

High frequency asymptotic scenario:

$$n \to \infty \qquad \Delta_n \to 0 \qquad n\Delta_n \to \infty$$

Condition 1: the process

- $\int_{x^\#}^{r} s(x;\theta)dx = \int_{\ell}^{x^\#} s(x;\theta)dx = \infty$ and $\int_{\ell}^{r} x^k \tilde{\mu}_\theta(x)dx = A(\theta) < \infty$

for all $k \in \mathbb{N}$, where $x^\#$ is an arbitrary point in (ℓ, r),

$s(x;\theta) = \exp\left(-2\int_{x^\#}^{x} \frac{b(y;\alpha)}{v(y;\beta)}dy\right)$ and $\tilde{\mu}_\theta(x) = [s(x;\theta)v(x;\beta)]^{-1}$

Condition 1: the process

- $\int_{x^{\#}}^{r} s(x;\theta)dx = \int_{\ell}^{x^{\#}} s(x;\theta)dx = \infty$ and $\int_{\ell}^{r} x^k \tilde{\mu}_\theta(x)dx = A(\theta) < \infty$

for all $k \in \mathbb{N}$, where $x^{\#}$ is an arbitrary point in (ℓ, r),

$s(x;\theta) = \exp\left(-2\int_{x^{\#}}^{x} \frac{b(y;\alpha)}{v(y;\beta)}dy\right)$ and $\tilde{\mu}_\theta(x) = [s(x;\theta)v(x;\beta)]^{-1}$

- $\sup_t E_\theta(|X_t|^k) < \infty$ for all $k \in \mathbb{N}$

Condition 1: the process

- $\int_{x^\#}^{r} s(x;\theta)dx = \int_{\ell}^{x^\#} s(x;\theta)dx = \infty$ and $\int_{\ell}^{r} x^k \tilde{\mu}_\theta(x)dx = A(\theta) < \infty$

for all $k \in \mathbb{N}$, where $x^\#$ is an arbitrary point in (ℓ, r),

$s(x;\theta) = \exp\left(-2\int_{x^\#}^{x} \frac{b(y;\alpha)}{v(y;\beta)}dy\right)$ and $\tilde{\mu}_\theta(x) = [s(x;\theta)v(x;\beta)]^{-1}$

- $\sup_t E_\theta(|X_t|^k) < \infty$ for all $k \in \mathbb{N}$

- $b, \sigma \in C_{p,4,1}((\ell, r) \times \Theta)$

Technical condition

$C_{p,k_1,k_2,k_3}(\mathbb{R}_+ \times (\ell, r)^2 \times \Theta)$ is the class of real functions $f(t, y, x; \theta)$ satisfying that

- $f(t, y, x; \theta)$ is k_1 times continuously differentiable with respect t, k_2 times continuously differentiable with respect y, and k_3 times continuously differentiable with respect α and with respect to β

- f and all partial derivatives $\partial_t^{i_1} \partial_y^{i_2} \partial_\alpha^{i_3} \partial_\beta^{i_4} f$, $i_j = 1, \dots k_j$, $j = 1, 2$, $i_3 + i_4 \leq k_3$, are of polynomial growth in x and y uniformly for θ in a compact set (for fixed t)

Technical condition

$C_{p,k_1,k_2,k_3}(\mathbb{R}_+ \times (\ell, r)^2 \times \Theta)$ is the class of real functions $f(t, y, x; \theta)$ satisfying that

- $f(t, y, x; \theta)$ is k_1 times continuously differentiable with respect t, k_2 times continuously differentiable with respect y, and k_3 times continuously differentiable with respect α and with respect to β

- f and all partial derivatives $\partial_t^{i_1} \partial_y^{i_2} \partial_\alpha^{i_3} \partial_\beta^{i_4} f$, $i_j = 1, \ldots k_j$, $j = 1, 2,$ $i_3 + i_4 \leq k_3$, are of polynomial growth in x and y uniformly for θ in a compact set (for fixed t)

$C_{p,k_1,k_2}((\ell, r) \times \Theta)$ for $f(y; \theta)$ and $C_{p,k_1,k_2}((\ell, r)^2 \times \Theta)$ for $f(y, x; \theta)$

are defined similarly

Technical condition

$C_{p,k_1,k_2,k_3}(\mathbb{R}_+ \times (\ell, r)^2 \times \Theta)$ is the class of real functions $f(t, y, x; \theta)$ satisfying that

- $f(t, y, x; \theta)$ is k_1 times continuously differentiable with respect t, k_2 times continuously differentiable with respect y, and k_3 times continuously differentiable with respect α and with respect to β

- f and all partial derivatives $\partial_t^{i_1} \partial_y^{i_2} \partial_\alpha^{i_3} \partial_\beta^{i_4} f$, $i_j = 1, \ldots k_j$, $j = 1, 2$, $i_3 + i_4 \leq k_3$, are of polynomial growth in x and y uniformly for θ in a compact set (for fixed t)

$C_{p,k_1,k_2}((\ell, r) \times \Theta)$ for $f(y; \theta)$ and $C_{p,k_1,k_2}((\ell, r)^2 \times \Theta)$ for $f(y, x; \theta)$

are defined similarly

$$|R(\Delta, y, x; \theta)| \leq F(y, x; \theta)$$

F is of polynomial growth in y and x uniformly for θ in a compact set

Condition 2: the estimating function

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_n, X_{t_i^n}, X_{t_{i-1}^n}; \theta) \qquad 2 - \text{dimensional}$$

- For some $\kappa \geq 2$

$$E_\theta(g(\Delta_n, X_{t_i^n}, X_{t_{i-1}^n}; \theta) \mid X_{t_{i-1}^n}) = \Delta_n^\kappa R(\Delta_n, X_{t_{i-1}^n}; \theta) \quad \text{for all } \theta \in \Theta$$

Condition 2: the estimating function

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_n, X_{t_i^n}, X_{t_{i-1}^n}; \theta) \qquad 2 - \text{dimensional}$$

- For some $\kappa \geq 2$

$$E_\theta(g(\Delta_n, X_{t_i^n}, X_{t_{i-1}^n}; \theta) \mid X_{t_{i-1}^n}) = \Delta_n^\kappa R(\Delta_n, X_{t_{i-1}^n}; \theta) \quad \text{for all } \theta \in \Theta$$

- The function $g(\Delta, y, x; \theta)$ has an expansion in powers of Δ

$$g(\Delta, y, x; \theta) = g(0, y, x; \theta) + \Delta g^{(1)}(y, x; \theta) + \frac{1}{2}\Delta^2 g^{(2)}(y, x; \theta) + \Delta^3 R(\Delta, y, x; \theta)$$

Condition 2: the estimating function

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_n, X_{t_i^n}, X_{t_{i-1}^n}; \theta) \qquad 2-\text{dimensional}$$

- For some $\kappa \geq 2$

$$E_\theta(g(\Delta_n, X_{t_i^n}, X_{t_{i-1}^n}; \theta) \mid X_{t_{i-1}^n}) = \Delta_n^\kappa R(\Delta_n, X_{t_{i-1}^n}; \theta) \quad \text{for all } \theta \in \Theta$$

- The function $g(\Delta, y, x; \theta)$ has an expansion in powers of Δ

$$g(\Delta, y, x; \theta) = g(0, y, x; \theta) + \Delta g^{(1)}(y, x; \theta) + \frac{1}{2}\Delta^2 g^{(2)}(y, x; \theta) + \Delta^3 R(\Delta, y, x; \theta)$$

- The function $R(\Delta, y, x; \theta)$ in the expansion of g is differentiable with respect to θ, and

$$\begin{aligned} g(\Delta, y, x; \theta) &\in C_{p,6,2}((\ell, r)^2 \times \Theta) \text{ for fixed } \Delta, \\ g^{(1)}(y, x; \theta) &\in C_{p,4,2}((\ell, r)^2 \times \Theta), \\ g^{(2)}(y, x; \theta) &\in C_{p,2,2}((\ell, r)^2 \times \Theta) \end{aligned}$$

Theorem 1

Suppose

- Conditions 1 and 2

Theorem 1

Suppose

- Conditions 1 and 2
- The identifiability condition that

$$\gamma(\theta,\theta_0) = \int_\ell^r [b(x,\alpha_0) - b(x,\alpha)]\partial_y g(0,x,x;\theta)\mu_{\theta_0}(x)dx + \frac{1}{2}\int_\ell^r [v(x,\beta_0) - v(x,\beta)]\partial_y^2 g(0,x,x;\theta)\mu_{\theta_0}(x)dx \neq 0$$

for all $\theta \neq \theta_0$

Theorem 1

Suppose

- Conditions 1 and 2
- The identifiability condition that

$$\gamma(\theta,\theta_0) = \int_\ell^r [b(x,\alpha_0) - b(x,\alpha)]\partial_y g(0,x,x;\theta)\mu_{\theta_0}(x)dx + \frac{1}{2}\int_\ell^r [v(x,\beta_0) - v(x,\beta)]\partial_y^2 g(0,x,x;\theta)\mu_{\theta_0}(x)dx \neq 0$$

for all $\theta \neq \theta_0$

- The matrix $S = \int_\ell^r A_{\theta_0}(x)\mu_{\theta_0}(x)dx$ is invertible, where

$$A_\theta(x) = \begin{pmatrix} \partial_\alpha b(x;\alpha)\partial_y g_1(0,x,x;\theta) & \frac{1}{2}\partial_\beta v(x;\beta)\partial_y^2 g_1(0,x,x;\theta) \\ \\ \partial_\alpha b(x;\alpha)\partial_y g_2(0,x,x;\theta) & \frac{1}{2}\partial_\beta v(x;\beta)\partial_y^2 g_2(0,x,x;\theta) \end{pmatrix}$$

Theorem 1

Then a consistent estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$ that solves the estimating equation $G_n(\theta) = 0$ exists and is unique in any compact subset of Θ containing θ_0 with a probability that goes to one as $n \to \infty$.

Theorem 1

Then a consistent estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$ that solves the estimating equation $G_n(\theta) = 0$ exists and is unique in any compact subset of Θ containing θ_0 with a probability that goes to one as $n \to \infty$.

For a martingale estimating function or more generally if $n\Delta^{2\kappa-1} \to 0$,

$$\sqrt{n\Delta_n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N_2\left(0, S^{-1}V_0(S^T)^{-1}\right)$$

under P_{θ_0}, where

$$V_0 = \int_\ell^r v(x, \beta_0)\partial_y g(0, x, x; \theta_0)\partial_y g(0, x, x; \theta_0)^T \mu_{\theta_0}(x)dx.$$

Optimal rate

Gobet (2002):

A discretely sampled diffusion is LAN in the high frequency asymptotics considered here, and the optimal rate of convergence is

For parameters in the drift coefficient: $\sqrt{n\Delta_n}$

For parameters in the diffusion coefficient: $\sqrt{n}$

Optimal rate

Gobet (2002):

A discretely sampled diffusion is LAN in the high frequency asymptotics considered here, and the optimal rate of convergence is

For parameters in the drift coefficient: $\sqrt{n\Delta_n}$

For parameters in the diffusion coefficient: $\sqrt{n}$

Jacobsen's condition:

$$\partial_y g_2(0, x, x; \theta) = 0$$

for all $x \in (\ell, r)$ and $\theta \in \Theta$

Optimal rate

Gobet (2002):

A discretely sampled diffusion is LAN in the high frequency asymptotics considered here, and the optimal rate of convergence is

For parameters in the drift coefficient: $\sqrt{n\Delta_n}$

For parameters in the diffusion coefficient: $\sqrt{n}$

Jacobsen's condition:

$$\partial_y g_2(0, x, x; \theta) = 0$$

for all $x \in (\ell, r)$ and $\theta \in \Theta$

Jacobsen (2001): small Δ-optimality

Theorem 2

Suppose

- Conditions 1 and 2

Theorem 2

Suppose

- Conditions 1 and 2
- The identifiability condition that

$$\int_\ell^r [b(x,\alpha_0) - b(x,\alpha)]\partial_y g_1(0,x,x;\theta)\mu_{\theta_0}(x)dx \neq 0 \qquad \text{when } \alpha \neq \alpha_0$$

$$\int_\ell^r [v(x,\beta_0) - v(x,\beta)]\partial_y^2 g_2(0,x,x;\theta)\mu_{\theta_0}(x)dx \neq 0 \qquad \text{when } \beta \neq \beta_0$$

Theorem 2

Suppose

- Conditions 1 and 2
- The identifiability condition that

$$\int_\ell^r [b(x,\alpha_0) - b(x,\alpha)]\partial_y g_1(0,x,x;\theta)\mu_{\theta_0}(x)dx \quad \neq \quad 0 \qquad \text{when } \alpha \neq \alpha_0$$

$$\int_\ell^r [v(x,\beta_0) - v(x,\beta)]\partial_y^2 g_2(0,x,x;\theta)\mu_{\theta_0}(x)dx \quad \neq \quad 0 \qquad \text{when } \beta \neq \beta_0$$

- $S_{11} \neq 0$ and $S_{22} \neq 0$

Theorem 2

Suppose

- Conditions 1 and 2
- The identifiability condition that

$$\int_\ell^r [b(x,\alpha_0) - b(x,\alpha)]\partial_y g_1(0,x,x;\theta)\mu_{\theta_0}(x)dx \quad \neq \quad 0 \qquad \text{when } \alpha \neq \alpha_0$$

$$\int_\ell^r [v(x,\beta_0) - v(x,\beta)]\partial_y^2 g_2(0,x,x;\theta)\mu_{\theta_0}(x)dx \quad \neq \quad 0 \qquad \text{when } \beta \neq \beta_0$$

- $S_{11} \neq 0$ and $S_{22} \neq 0$
- $\partial_y g_2(0,x,x;\theta) = 0$

Theorem 2

Then a consistent estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$ that solves the estimating equation $G_n(\theta) = 0$ exists and is unique in any compact subset of Θ containing θ_0 with a probability that goes to one as $n \to \infty$.

Theorem 2

Then a consistent estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$ that solves the estimating equation $G_n(\theta) = 0$ exists and is unique in any compact subset of Θ containing θ_0 with a probability that goes to one as $n \to \infty$.

If, moreover,

$$\partial_\alpha \partial_y^2 g_2(0, x, x; \theta) = 0,$$

then for a martingale estimating function or more generally if $n\Delta^{2(\kappa-1)} \to 0$,

$$\begin{pmatrix} \sqrt{n\Delta_n}(\hat{\alpha}_n - \alpha_0) \\ \sqrt{n}(\hat{\beta}_n - \beta_0) \end{pmatrix} \xrightarrow{\mathcal{D}} N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} W_1/S_{11}^2 & 0 \\ 0 & W_2/S_{22}^2 \end{pmatrix} \right)$$

where

$$W_1 = \int_\ell^r v(x;\beta_0)[\partial_y g_1(0,x,x;\theta_0)]^2 \mu_{\theta_0}(x)dx$$

and

$$W_2 = \frac{1}{2}\int_\ell^r v(x;\beta_0)^2[\partial_y^2 g_2(0,x,x;\theta_0)]^2 \mu_{\theta_0}(x)dx$$

Efficiency - 1

Gobet (2002):

A discretely sampled diffusion is LAN in the high frequency asymptotics considered here, and the Fisher information is

$$\begin{pmatrix} \int_\ell^r \frac{(\partial_\alpha b(x;\alpha_0))^2}{v(x;\beta_0)} \mu_{\theta_0}(x)dx & 0 \\ 0 & \frac{1}{2}\int_\ell^r \left[\frac{\partial_\beta v(x;\beta_0)}{v(x;\beta_0)}\right]^2 \mu_{\theta_0}(x)dx \end{pmatrix}$$

Efficiency - 1

Gobet (2002):

A discretely sampled diffusion is LAN in the high frequency asymptotics considered here, and the Fisher information is

$$\begin{pmatrix} \int_\ell^r \frac{(\partial_\alpha b(x;\alpha_0))^2}{v(x;\beta_0)} \mu_{\theta_0}(x)dx & 0 \\ 0 & \frac{1}{2}\int_\ell^r \left[\frac{\partial_\beta v(x;\beta_0)}{v(x;\beta_0)}\right]^2 \mu_{\theta_0}(x)dx \end{pmatrix}$$

Condition for efficiency:

$$\partial_y g_1(0,x,x;\theta) = \partial_\alpha b(x;\alpha)/v(x;\beta) \qquad \partial_y^2 g_2(0,x,x;\theta) = \partial_\beta v(x;\beta)/v(x;\beta)^2$$

for all $x \in (\ell, r)$ and $\theta \in \Theta$

Efficiency - 1

Gobet (2002):

A discretely sampled diffusion is LAN in the high frequency asymptotics considered here, and the Fisher information is

$$\begin{pmatrix} \int_\ell^r \frac{(\partial_\alpha b(x;\alpha_0))^2}{v(x;\beta_0)} \mu_{\theta_0}(x)dx & 0 \\ 0 & \frac{1}{2}\int_\ell^r \left[\frac{\partial_\beta v(x;\beta_0)}{v(x;\beta_0)}\right]^2 \mu_{\theta_0}(x)dx \end{pmatrix}$$

Condition for efficiency:

$$\partial_y g_1(0,x,x;\theta) = \partial_\alpha b(x;\alpha)/v(x;\beta) \qquad \partial_y^2 g_2(0,x,x;\theta) = \partial_\beta v(x;\beta)/v(x;\beta)^2$$

for all $x \in (\ell, r)$ and $\theta \in \Theta$

Jacobsen (2001): small Δ-optimality

Quadratic martingale estimating functions

$$\sum_{i=1}^{n} \begin{pmatrix} a_1(X_{t_{i-1}^n}, \Delta; \theta)(X_{t_i^n} - F(\Delta, X_{t_{i-1}^n}; \theta)) \\ \\ a_2(X_{t_{i-1}^n}, \Delta; \theta) \left[(X_{t_i^n} - F(\Delta, X_{t_{i-1}^n}; \theta))^2 - \phi(\Delta, X_{t_{i-1}^n}; \theta) \right] \end{pmatrix}$$

$$\begin{aligned} F(\Delta, x; \theta) &= E_\theta(X_\Delta | X_0 = x) = x + O(\Delta) \\ \phi(\Delta, x; \theta) &= \text{Var}_\theta(X_\Delta | X_0 = x) = O(\Delta) \end{aligned}$$

Quadratic martingale estimating functions

$$\sum_{i=1}^{n} \begin{pmatrix} a_1(X_{t_{i-1}^n}, \Delta; \theta)(X_{t_i^n} - F(\Delta, X_{t_{i-1}^n}; \theta)) \\ \\ a_2(X_{t_{i-1}^n}, \Delta; \theta)\left[(X_{t_i^n} - F(\Delta, X_{t_{i-1}^n}; \theta))^2 - \phi(\Delta, X_{t_{i-1}^n}; \theta)\right] \end{pmatrix}$$

$$\begin{aligned} F(\Delta, x; \theta) &= E_\theta(X_\Delta | X_0 = x) = x + O(\Delta) \\ \phi(\Delta, x; \theta) &= \mathrm{Var}_\theta(X_\Delta | X_0 = x) = O(\Delta) \end{aligned}$$

$$g(0, y, x; \theta) = \begin{pmatrix} a_1(x, 0; \theta)(y - x) \\ a_2(x, 0; \theta)(y - x)^2 \end{pmatrix}$$

Quadratic martingale estimating functions

$$\sum_{i=1}^{n} \begin{pmatrix} a_1(X_{t_{i-1}^n}, \Delta; \theta)(X_{t_i^n} - F(\Delta, X_{t_{i-1}^n}; \theta)) \\ \\ a_2(X_{t_{i-1}^n}, \Delta; \theta) \left[(X_{t_i^n} - F(\Delta, X_{t_{i-1}^n}; \theta))^2 - \phi(\Delta, X_{t_{i-1}^n}; \theta) \right] \end{pmatrix}$$

$$\begin{aligned} F(\Delta, x; \theta) &= E_\theta(X_\Delta | X_0 = x) = x + O(\Delta) \\ \phi(\Delta, x; \theta) &= \text{Var}_\theta(X_\Delta | X_0 = x) = O(\Delta) \end{aligned}$$

$$g(0, y, x; \theta) = \begin{pmatrix} a_1(x, 0; \theta)(y - x) \\ a_2(x, 0; \theta)(y - x)^2 \end{pmatrix}$$

$$\begin{aligned} \partial_y g_2(0, y, x; \theta) &= 2a_2(x, \Delta; \theta)(y - x) \quad \text{Jacobsen's condition satisfied} \\ \partial_y g_1(0, x, x; \theta) &= a_1(x, 0; \theta) = \partial_\alpha b(x; \alpha)/v(x; \beta) \\ \partial_y^2 g_2(0, x, x; \theta) &= 2a_2(x, 0; \theta) = \partial_\beta v(x; \beta)/v(x; \beta)^2 \end{aligned}$$

Martingale estimating functions

$$
\begin{aligned}
g(\Delta, y, x; \theta) &= \sum_{j=1}^{N} a_j(x, \Delta; \theta)[f_j(y; \theta) - \pi_\theta^\Delta f_j(x; \theta)] \\
&= A(x, \Delta_n; \theta)[f(y; \theta) - \pi_\theta^{\Delta_n} f(x; \theta)]
\end{aligned}
$$

$$
G_n(\theta) = \sum_{i=1}^{n} A(X_{t_{i-1}^n}, \Delta_n; \theta)[f(X_{t_i^n}; \theta) - \pi_\theta^{\Delta_n} f(X_{t_{i-1}^n}; \theta)]
$$

$f(y; \theta) = (f_1(y; \theta), \ldots, f_N(y; \theta))^T$

$A(x, \Delta; \theta)$ a $2 \times N$-matrix of weights

$\pi_\theta^\Delta f(x; \theta) = E_\theta(f(X_\Delta; \theta) \mid X_0 = x)$ is the transition operator

Efficiency - 2

Suppose Condition 1 is satisfied and that the functions f_j are twice continuously differentiable.

A sufficient condition that it is possible to find a specification of the weight matrix $A(x, \Delta; \theta)$ such that the estimating function $G_n(\theta)$ gives estimators that are rate optimal and efficient is that

Efficiency - 2

Suppose Condition 1 is satisfied and that the functions f_j are twice continuously differentiable.

A sufficient condition that it is possible to find a specification of the weight matrix $A(x, \Delta; \theta)$ such that the estimating function $G_n(\theta)$ gives estimators that are rate optimal and efficient is that

- $N \geq 2$
- and that the matrix

$$D(x) = \begin{pmatrix} f_1'(x) & f_1''(x) \\ \\ f_2'(x) & f_2''(x) \end{pmatrix}$$

is invertible for μ_θ-almost all x.

Efficiency - 2

Suppose Condition 1 is satisfied and that the functions f_j are twice continuously differentiable.

A sufficient condition that it is possible to find a specification of the weight matrix $A(x, \Delta; \theta)$ such that the estimating function $G_n(\theta)$ gives estimators that are rate optimal and efficient is that

- $N \geq 2$
- and that the matrix

$$D(x) = \begin{pmatrix} f_1'(x) & f_1''(x) \\ & \\ f_2'(x) & f_2''(x) \end{pmatrix}$$

is invertible for μ_θ-almost all x.

Jacobsen (2002)

For a d-dimensional diffusion: $N \geq d(d+3)/2$

Godambe-Heyde optimality

$$G_n(\theta) = \sum_{i=1}^{n} A^*(X_{t_{i-1}^n}, \Delta_n; \theta)[f(X_{t_i^n}) - \pi_\theta^{\Delta_n} f(X_{t_{i-1}^n})]$$

is Godambe-Heyde optimal if

$$A^*(x, \Delta; \theta) E_\theta \left([f(X_\Delta) - \pi_\theta^\Delta f(x)][f(X_\Delta) - \pi_\theta^\Delta f(x)]^T \mid X_0 = x \right) = \partial_\theta \pi_\theta^\Delta f^T(x)$$

for μ_θ-almost all x

Efficiency - 3

Suppose Condition 1 is satisfied, that the functions f_j are six times continuously differentiable, that $N \geq 2$ and that $D(x)$ is invertible for μ_θ-almost all x.

Then

$$g^*(\Delta, y, x; \theta) = \begin{pmatrix} 1 & 0 \\ & \\ 0 & 2\Delta \end{pmatrix} A^*(x, \Delta; \theta)[f(y) - \pi_\theta^\Delta f(x)]$$

satisfies that

$$\partial_y g_2^*(0, x, x; \theta) = 0$$

and

$$\partial_y g_1^*(0, x, x; \theta) = \partial_\alpha b(x; \alpha)/v(x; \beta) \qquad \partial_y^2 g_2^*(0, x, x; \theta) = \partial_\beta v(x; \beta)/v(x; \beta)^2$$

for all $x \in (\ell, r)$ and $\theta \in \Theta$

Asymptotic scenarios - 1

DATA: $X_0, X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$

Possibly time-series from several individuals

Asymptotic scenarios - 1

DATA: $X_0, X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$

Possibly time-series from several individuals

LARGE SAMPLE ASYMPTOTICS:

$n \to \infty$ or

number of individuals $\to \infty$ (Pedersen, 2000)

Asymptotic scenarios - 1

DATA: $X_0, X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$

Possibly time-series from several individuals

LARGE SAMPLE ASYMPTOTICS:

$n \to \infty$ or

number of individuals $\to \infty$ (Pedersen, 2000)

HIGH FREQUENCY ASYMPTOTICS:

$\Delta \to 0$ and $n \to \infty$

$n\Delta \to \infty$: Prakasa-Rao (1983), Yoshida (1992), Kessler (1997), Sørensen (2007)

$n\Delta$ constant: Dohnal (1987), Jacod and Genon-Catalot (1993)

Asymptotic scenarios - 2

SMALL DIFFUSION ASYMPTOTICS:

$\sigma(X;\theta) = \epsilon g(x,\theta) \qquad \epsilon \to 0$ Sørensen (2000)

Asymptotic scenarios - 2

SMALL DIFFUSION ASYMPTOTICS:

$\sigma(X;\theta) = \epsilon g(x,\theta) \qquad \epsilon \to 0$ Sørensen (2000)

SMALL DIFFUSION/HIGH FREQUENCY ASYMPTOTICS:

$\sigma(X;\theta) = \epsilon g(x,\theta) \qquad \epsilon \to 0$, $\Delta \to 0$ and $n \to \infty$

Genon-Catalot (1990), Sørensen and Uchida (2003), Gloter and Sørensen (2006)

Asymptotic scenarios - 2

SMALL DIFFUSION ASYMPTOTICS:

$\sigma(X;\theta) = \epsilon g(x,\theta) \qquad \epsilon \to 0$ Sørensen (2000)

SMALL DIFFUSION/HIGH FREQUENCY ASYMPTOTICS:

$\sigma(X;\theta) = \epsilon g(x,\theta) \qquad \epsilon \to 0$, $\Delta \to 0$ and $n \to \infty$

Genon-Catalot (1990), Sørensen and Uchida (2003), Gloter and Sørensen (2006)

SMALL VOLATILITY OF VOLATILITY ASYMPTOTICS

Sørensen and Yoshida (2000)

Explicit martingale estimating functions

Kessler and Sørensen (1999)

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

Generator:

$$L_\theta = \tfrac{1}{2}\sigma^2(x;\theta)\frac{d^2}{dx^2} + b(x;\theta)\frac{d}{dx},$$

φ eigenfunction for L_θ:

$$L_\theta\varphi = -\lambda_\theta\varphi$$

Explicit martingale estimating functions

Kessler and Sørensen (1999)

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

Generator:

$$L_\theta = \tfrac{1}{2}\sigma^2(x;\theta)\frac{d^2}{dx^2} + b(x;\theta)\frac{d}{dx},$$

φ eigenfunction for L_θ:

$$L_\theta\varphi = -\lambda_\theta\varphi$$

Under weak regularity conditions

$$\pi_\theta^\Delta\varphi(x) = E_\theta(\varphi(X_\Delta)|X_0 = x) = e^{-\lambda_\theta\Delta}\varphi(x)$$

i.e. φ is an eigenfunction for π_θ^Δ

Explicit martingale estimating functions

Three sets of sufficient conditions ensuring that $\pi_\theta^\Delta \varphi(x) = e^{-\lambda_\theta \Delta}\varphi(x)$:

(i) X ergodic with invariant measure μ, and $\int \varphi'(x)^2 \sigma^2(x)\mu(dx) < \infty$

(ii) σ and φ' bounded

(iii) b and σ of linear growth, φ' of polynomial growth

Pearson diffusions

Wong (1964), Zhou (2003), Forman & Sørensen (2006)

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta(aX_t^2 + bX_t + c)}dW_t, \qquad \beta > 0$$

$$L\varphi = \beta(ax^2 + bx + c)\varphi'' + \beta(x - \mu)\varphi'$$

Pearson diffusions

Wong (1964), Zhou (2003), Forman & Sørensen (2006)

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta(aX_t^2 + bX_t + c)}dW_t, \qquad \beta > 0$$

$$L\varphi = \beta(ax^2 + bx + c)\varphi'' + \beta(x - \mu)\varphi'$$

If φ is a polynomial of order k, then so is $L\varphi$

Thus we can find eigenfunctions that are explicit polynomials

Pearson diffusions

Wong (1964), Zhou (2003), Forman & Sørensen (2006)

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta(aX_t^2 + bX_t + c)}dW_t, \qquad \beta > 0$$

$$L\varphi = \beta(ax^2 + bx + c)\varphi'' + \beta(x - \mu)\varphi'$$

If φ is a polynomial of order k, then so is $L\varphi$

Thus we can find eigenfunctions that are explicit polynomials

$$\varphi_n(x) = \sum_{j=0}^{n} p_{n,j} x^j, \qquad p_{n,n} = 1$$

$$(a_j - a_n)p_{n,j} = b_{j+1}p_{n,j+1} + c_{j+2}p_{n,j+2}, \quad j = 0, \ldots, n-1, \quad p_{n,n+1} = 0$$

$$a_j = j\{1 - (j-1)a\}\beta, \quad b_j = j\{\mu + (j-1)b\}\beta, \quad c_j = j(j-1)c\beta$$

Pearson diffusions

Eigenvalues: $\lambda_n = a_n$

$a < (n-1)^{-1}$: φ_n integrable w.r.t. the invariant distribution

$a < (2n-1)^{-1}$: φ_n square integrable w.r.t. the invariant distribution

Pearson diffusions

Eigenvalues: $\lambda_n = a_n$

$a < (n-1)^{-1}$: φ_n integrable w.r.t. the invariant distribution

$a < (2n-1)^{-1}$: φ_n square integrable w.r.t. the invariant distribution

The class of possible stationary marginal distributions is equal to Pearson's system of distributions. Up to location-scale transformations the following is a complete list

Pearson diffusions

Eigenvalues: $\lambda_n = a_n$

$a < (n-1)^{-1}$: φ_n integrable w.r.t. the invariant distribution

$a < (2n-1)^{-1}$: φ_n square integrable w.r.t. the invariant distribution

The class of possible stationary marginal distributions is equal to Pearson's system of distributions. Up to location-scale transformations the following is a complete list

- Normal distribution:
  Ornstein-Uhlenbeck process: $\sigma^2(x) = 2\beta c$, $c > 0$
  State space: the real line
  Hermite polynomials

Pearson diffusions

- Gamma distribution:

  CIR process: $\sigma^2(x) = 2\beta bx$, $b > 0$

  State space: the positive real axis

  Laguerre polynomials

Pearson diffusions

- Gamma distribution:
  CIR process: $\sigma^2(x) = 2\beta bx$, $b > 0$
  State space: the positive real axis
  Laguerre polynomials
- Beta distribution:
  Jacobi diffusions: $\sigma^2(x) = -2\beta ax(1-x)$, $a < 0$
  State space: the interval $(0, 1)$
  Jacobi polynomials

Pearson diffusions

- Gamma distribution:
  CIR process: $\sigma^2(x) = 2\beta bx$, $b > 0$
  State space: the positive real axis
  Laguerre polynomials

- Beta distribution:
  Jacobi diffusions: $\sigma^2(x) = -2\beta ax(1-x)$, $a < 0$
  State space: the interval $(0, 1)$
  Jacobi polynomials

- Inverse gamma distribution:
  "GARCH"-diffusions: $\sigma^2(x) = 2\beta ax^2$, $a > 0$
  State space: the positive real axis
  Bessel polynomials

Pearson diffusions

- Gamma distribution:
  CIR process: $\sigma^2(x) = 2\beta bx$, $b > 0$
  State space: the positive real axis
  Laguerre polynomials

- Beta distribution:
  Jacobi diffusions: $\sigma^2(x) = -2\beta ax(1-x)$, $a < 0$
  State space: the interval $(0, 1)$
  Jacobi polynomials

- Inverse gamma distribution:
  "GARCH"-diffusions: $\sigma^2(x) = 2\beta ax^2$, $a > 0$
  State space: the positive real axis
  Bessel polynomials

- F-distribution:
  $\sigma^2(x) = 2\beta ax(x+1)$
  State space: the positive real axis

Pearson diffusions

- t-distribution with $\nu = 1 + 1/a$ degrees of freedom:

$$\mu = 0 \text{ and } \sigma^2(x) = 2\beta a(x^2 + 1), \quad a > 0$$

State space: the real line

Pearson diffusions

- t-distribution with $\nu = 1 + 1/a$ degrees of freedom:

$$\mu = 0 \text{ and } \sigma^2(x) = 2\beta a(x^2 + 1), \quad a > 0$$

State space: the real line

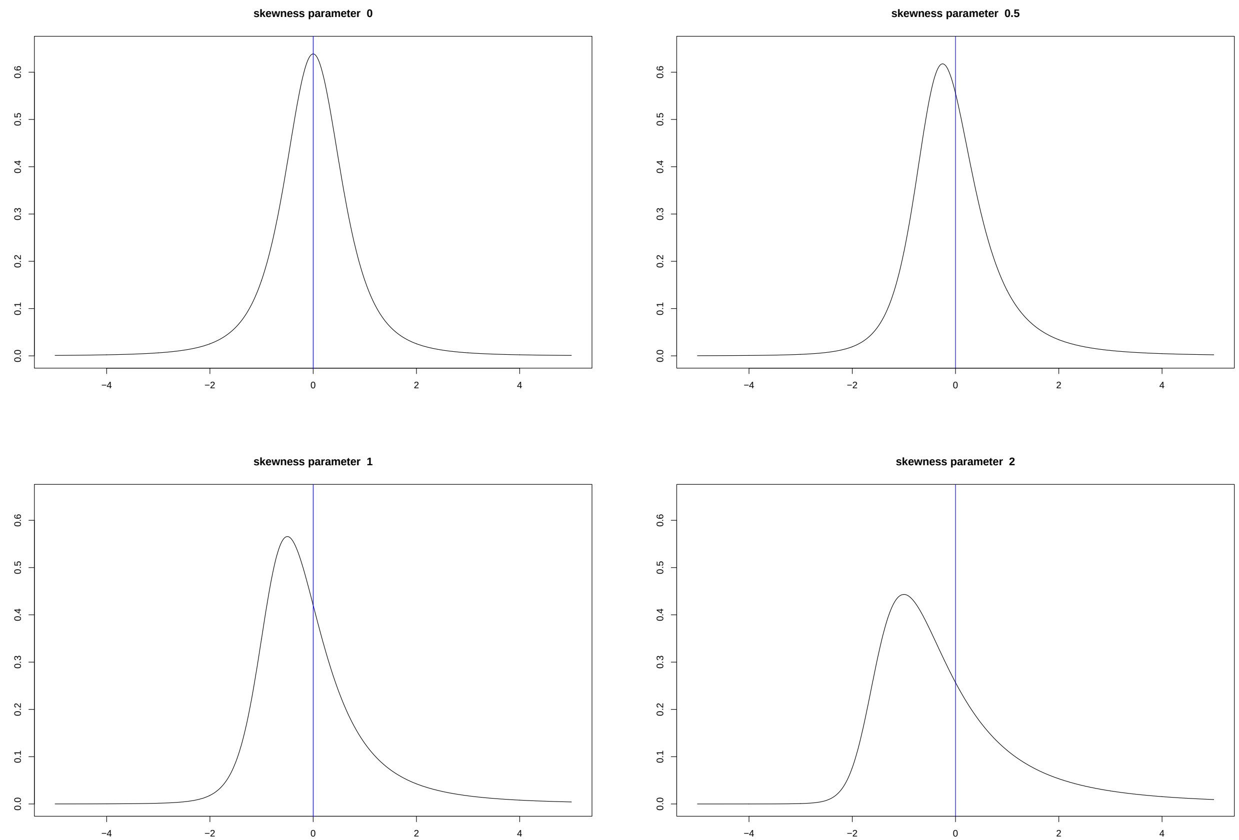
- Pearson's type IV distribution, a skew t-distribution

$$dZ_t = -\beta Z_t dt + \sqrt{2\beta(\nu - 1)^{-1}\{Z_t^2 + 2\rho\nu^{\frac{1}{2}} Z_t + (1 + \rho^2)\nu\}} dW_t$$

$$f(z) \propto \{(z/\sqrt{\nu} + \rho)^2 + 1\}^{-(\nu+1)/2} \exp\left\{\rho(\nu + 1)\tan^{-1}\left(z/\sqrt{\nu} + \rho\right)\right\}$$

An expression for the normalizing constant when ν is an integer can be found in Nagahara (1996), who used this diffusion to model the Nikkei 225 index, the TOPIX index and the Standard and Poors 500 index

Pearson's type IV distribution



Densities of skew t-distributions (Pearson's type IV distributions) with zero mean for $\rho = 0$, 0.5, 1, and 2 respectively

Transformations of Pearson diffusions

X_t: $\varphi(x)$ eigenfunction with eigenvalue λ

$T(X_t)$: $\varphi(T^{-1}(x))$ eigenfunction with eigenvalue λ T an injection

Transformations of Pearson diffusions

X_t: $\varphi(x)$ eigenfunction with eigenvalue λ

$T(X_t)$: $\varphi(T^{-1}(x))$ eigenfunction with eigenvalue λ T an injection

Jacobi diffusion state space $(-1, 1)$, $\beta, \sigma > 0$, $\gamma \in (-1, 1)$

$$dX_t = -\beta[X_t - \gamma]dt + \sigma\sqrt{1 - X_t^2}dW_t$$

Eigenfunctions: $P_n^{(\beta(1-\gamma)\sigma^{-2}-1,\, \beta(1+\gamma)\sigma^{-2}-1)}(x)$

$P_n^{(a,b)}(x)$ denotes the Jacobi polynomial of order n

Transformations of Pearson diffusions

X_t: $\varphi(x)$ eigenfunction with eigenvalue λ

$T(X_t)$: $\varphi(T^{-1}(x))$ eigenfunction with eigenvalue λ T an injection

Jacobi diffusion state space $(-1,1)$, $\beta,\sigma > 0$, $\gamma \in (-1,1)$

$$dX_t = -\beta[X_t - \gamma]dt + \sigma\sqrt{1 - X_t^2}dW_t$$

Eigenfunctions: $P_n^{(\beta(1-\gamma)\sigma^{-2}-1,\,\beta(1+\gamma)\sigma^{-2}-1)}(x)$

$P_n^{(a,b)}(x)$ denotes the Jacobi polynomial of order n

$Y_t = \sin^{-1}(X_t)$ state space $(-\frac{\pi}{2}, \frac{\pi}{2})$, $\rho = \beta - \frac{1}{2}\sigma^2$, $\varphi = \beta\gamma/(\beta - \frac{1}{2}\sigma^2)$

$$dY_t = -\rho\frac{\sin(Y_t) - \varphi}{\cos(Y_t)}dt + \sigma d\tilde{W}_t$$

Eigenfunctions: $P_n^{(\rho(1-\varphi)\sigma^{-2}-\frac{1}{2},\,\rho(1+\varphi)\sigma^{-2}-\frac{1}{2})}(\sin(x))$

Optimal martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta),$$

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta) h_j(\Delta, X_{t_i}, X_{t_{i-1}}; \theta)$$

$$G_n(\theta) = \sum_{i=1}^{n} A(X_{t_{i-1}}, \Delta_i; \theta) h(\Delta_i, X_{t_i}, X_{t_{i-1}}; \theta)$$

$h = (h_1, \ldots, h_N)^T$, $\qquad h_j(\Delta, x, y; \theta) = f_j(y) - \pi_\theta^\Delta f_j(x)$

$A(x, \Delta; \theta)$ $p \times N$-matrix of weights

Explicit optimal estimating functions

We consider

$$h_i(\Delta, x, y; \theta) = \varphi_i(y;\theta) - e^{-\lambda_i(\theta)\Delta}\varphi_i(x;\theta), \qquad i = 1, \ldots, N,$$

where $\varphi_i(y;\theta)$ is an eigenfunction of the generator with eigenvalue $\lambda_i(\theta)$, for which $\pi_\theta^\Delta \varphi_i(x;\theta) = e^{-\lambda_i(\theta)\Delta}\varphi_i(x;\theta)$

Suppose

$$\varphi_i(x;\theta) = \psi_i(\kappa(x);\theta),$$

where κ is a real function independent of θ, and ψ_i is a polynomial of degree i:

$$\psi_i(y;\theta) = \sum_{j=0}^{i} a_{i,j}(\theta)\, y^j$$

Explicit optimal martingale estimating functions

Optimal weight matrix:

$$A^*(x,\Delta;\theta) = B(x,\Delta;\theta)^T V(x,\Delta;\theta)^{-1}$$

$$B(x,\Delta;\theta)_{ij} = \sum_{k=0}^{j} \partial_{\theta_i} a_{j,k}(\theta)\pi_\theta^\Delta \kappa^k(x) - \partial_{\theta_i}[e^{-\lambda_j(\theta)\Delta}\varphi_j(x;\theta)]$$

$i = 1,\ldots,p, \;\; j = 1,\ldots,N$

$$V(\Delta,x;\theta)_{ij} = \sum_{r=0}^{i}\sum_{s=0}^{j} a_{i,r}(\theta)a_{j,s}(\theta)\pi_\theta^\Delta \kappa^{r+s}(x) - e^{-[\lambda_i(\theta)+\lambda_j(\theta)]\Delta}\,\varphi_i(x;\theta)\varphi_j(x;\theta)$$

$i, j = 1,\ldots,N$

Explicit optimal martingale estimating functions

Optimal weight matrix:

$$A^*(x,\Delta;\theta) = B(x,\Delta;\theta)^T V(x,\Delta;\theta)^{-1}$$

$$B(x,\Delta;\theta)_{ij} = \sum_{k=0}^{j} \partial_{\theta_i} a_{j,k}(\theta) \pi_\theta^\Delta \kappa^k(x) - \partial_{\theta_i}[e^{-\lambda_j(\theta)\Delta}\varphi_j(x;\theta)]$$

$i = 1, \ldots, p,\ \ j = 1, \ldots, N$

$$V(\Delta, x;\theta)_{ij} = \sum_{r=0}^{i} \sum_{s=0}^{j} a_{i,r}(\theta) a_{j,s}(\theta) \pi_\theta^\Delta \kappa^{r+s}(x) - e^{-[\lambda_i(\theta)+\lambda_j(\theta)]\Delta} \varphi_i(x;\theta)\varphi_j(x;\theta)$$

$i, j = 1, \ldots, N$

Explicit optimal estimating functions

Thus to find the optimal estimating function based on the first N eigenfunctions, we need to find the conditional moments

$$\pi_\theta^\Delta \kappa^i(x) = E_\theta(\kappa(X_\Delta)^i \,|\, X_0 = x) \quad \text{for} \quad 1 \leq i \leq 2N$$

If we apply the linear operator π_θ^Δ to both sides of

$$\varphi_i(y;\theta) = \sum_{j=0}^{i} a_{i,j}(\theta)\,\kappa(y)^j$$

for $i = 1, \ldots, 2N$, we obtain a system of linear equations

$$e^{-\lambda_i(\theta)\Delta}\varphi_i(x;\theta) = \sum_{j=0}^{i} a_{i,j}(\theta)\pi_\theta^\Delta \kappa^j(x), \qquad i = 1, \ldots, 2N$$