Jørlunde, January 2009

Estimation for stochastic differential equations: tractable models and efficiency

Michael Sørensen
University of Copenhagen, Denmark

http://www.math.ku.dk/∼michael

Discretely observed diffusion

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

Discretely observed diffusion

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

$$X_{(n+1)\delta} \approx X_{n\delta} + b(X_{n\delta};\theta)\delta + \sigma(X_{n\delta};\theta)Z_{n+1} \qquad Z_{n+1} \sim N(0,\delta), \text{ i.i.d.}$$

Discretely observed diffusion

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

$$X_{(n+1)\delta} \approx X_{n\delta} + b(X_{n\delta};\theta)\delta + \sigma(X_{n\delta};\theta)Z_{n+1} \qquad Z_{n+1} \sim N(0,\delta), \quad \text{i.i.d.}$$

$$\frac{dX_t}{dt} = b(X_t;\theta)$$

Discretely observed diffusion

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

$$X_{(n+1)\delta} \approx X_{n\delta} + b(X_{n\delta};\theta)\delta + \sigma(X_{n\delta};\theta)Z_{n+1} \qquad Z_{n+1} \sim N(0,\delta), \text{ i.i.d.}$$

$$\frac{dX_t}{dt} = b(X_t;\theta)$$

Data: $X_{t_0}, \cdots, X_{t_n} \qquad \Delta_i = t_i - t_{i-1}$

Discretely observed diffusion

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

$$X_{(n+1)\delta} \approx X_{n\delta} + b(X_{n\delta};\theta)\delta + \sigma(X_{n\delta};\theta)Z_{n+1} \qquad Z_{n+1} \sim N(0,\delta), \text{ i.i.d.}$$

$$\frac{dX_t}{dt} = b(X_t;\theta)$$

Data: $X_{t_0}, \cdots, X_{t_n} \qquad \Delta_i = t_i - t_{i-1}$

Review papers:

Helle Sørensen (2004) Int. Stat. Rev.

Bibby, Jacobsen and Sørensen (2004)

Sørensen (2008,2009)

Martingale estimating functions

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

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

$\uparrow$ $\uparrow$

p-dimensional real valued

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

Martingale estimating functions

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

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

$\uparrow$ p-dimensional $\uparrow$ 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

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

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

$\uparrow$ $\uparrow$

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

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:

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

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

Pearson diffusions

Wong (1964), Zhou (2003), Forman & Sørensen (2008)

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

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

Pearson diffusions

Wong (1964), Zhou (2003), Forman & Sørensen (2008)

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

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

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 (2008)

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

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

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

The class of possible stationary marginal distributions is equal to Pearson's system of distributions

$Y_t = aX_t + b$ is also a Pearson diffusion

Up to location-scale transformations the following is a complete list

Pearson diffusions

- Normal distribution:

  Ornstein-Uhlenbeck process:

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

  $X_t \sim N(\mu, c)$

  State space: the real line

  Eigenfunctions: Hermite polynomials

Pearson diffusions

- Normal distribution:

  Ornstein-Uhlenbeck process:

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

  $X_t \sim N(\mu, c)$

  State space: the real line

  Eigenfunctions: Hermite polynomials

- Gamma-distribution:

  Square root process:

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta b X_t}\, dW_t, \quad b > 0$$

  X_t gamma-distributed with mean μ and scale parameter b

  State space: the positive real axis

  Eigenfunctions: Laguerre polynomials

Pearson diffusions

- Beta-distribution:

  Jacobi diffusions:

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta a X_t(1 - X_t)}\, dW_t, \quad a > 0$$

  X_t beta-distributed with $p(x) \propto x^{\mu/a-1}(1-x)^{(1-\mu/a)-1}$

  State space: the interval $(0, 1)$

  Eigenfunctions: Jacobi polynomials

Pearson diffusions

- Beta-distribution:

  Jacobi diffusions:

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta a X_t(1 - X_t)}\, dW_t, \quad a > 0$$

  X_t beta-distributed with $p(x) \propto x^{\mu/a-1}(1-x)^{(1-\mu/a)-1}$

  State space: the interval $(0, 1)$

  Eigenfunctions: Jacobi polynomials

- Inverse gamma distribution:

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta a}\, X_t\, dW_t, \quad a > 0$$

  Density of X_t: $\quad p(x) \propto x^{-(a^{-1}+2)} \exp(-\frac{\mu}{ax})$

  State space: the positive real axis

  Eigenfunctions: Bessel polynomials

Pearson diffusions

- F-distribution:

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta a X_t(X_t + 1)}\, dW_t, \quad a > 0$$

$(1 + a)\mu^{-1}X_t$ F-distributed with $2\mu a^{-1}$ and $2a^{-1} + 2$ degrees of freedom
State space: the positive real axis

Pearson diffusions

- F-distribution:

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta a X_t(X_t + 1)}\, dW_t, \quad a > 0$$

$(1 + a)\mu^{-1}X_t$ F-distributed with $2\mu a^{-1}$ and $2a^{-1} + 2$ degrees of freedom
State space: the positive real axis

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

$$dX_t = -\beta(X_t - \mu)dt + \sqrt{2\beta a(X_t^2 + 1)}\, dW_t, \quad a > 0$$

X_t is t-distribution with $1 + 1/a$ degrees of freedom and mean μ
State space: the real line

Pearson diffusions

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

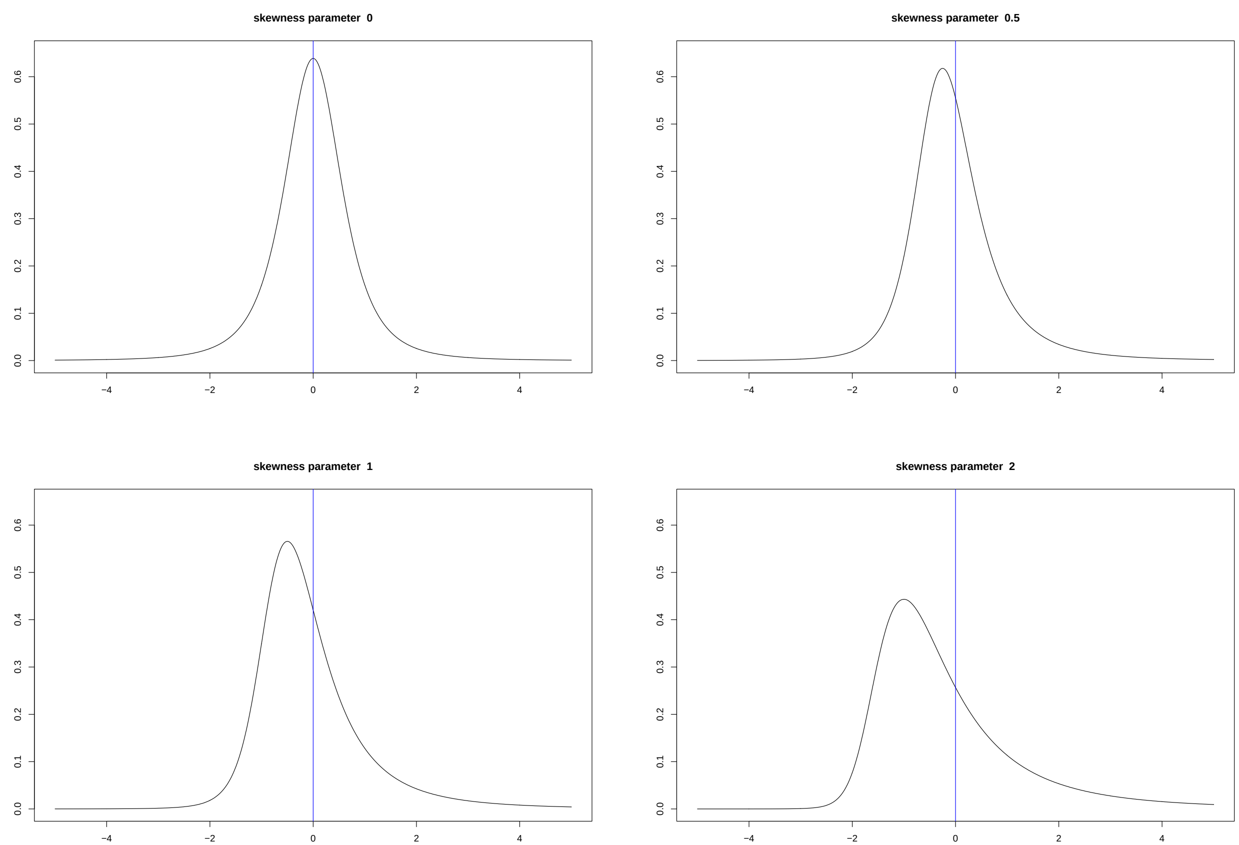
$$dZ_t = -\beta Z_t dt + \sqrt{2\beta(\nu-1)^{-1}\{Z_t^2 + 2\rho\nu^{\frac{1}{2}} Z_t + (1+\rho^2)\nu\}} dW_t, \quad \nu > 1$$

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

An expression for the normalizing constant when $\nu \in \mathbb{N}$ can be found in Nagahara (1996)

$\rho = 0$: t-distribution with ν degrees of freedom

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

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

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta) \left[\varphi_j(y;\theta) - e^{-\lambda_j(\theta)\Delta} \varphi_j(x;\theta)\right]$$

Optimal martingale estimating functions

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

$$g(\Delta, y, x; \theta) = \sum_{j=1}^{N} a_j(x, \Delta; \theta) \left[\varphi_j(y; \theta) - e^{-\lambda_j(\theta)\Delta}\varphi_j(x; \theta)\right]$$

Suppose

$$\varphi_j(x; \theta) = \psi_j(\kappa(x); \theta),$$

where κ is a real function independent of θ, and ψ_j is a polynomial of degree j:

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

Then the optimal weights $a_j^*(x, \Delta; \theta)$ can be found explicitly

Asymptotics - low frequency

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

Assume that X is ergodic with invariant measure $\mu_\theta(x)$, that $t_i = \Delta i$, 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}. Here

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

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

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

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

$\theta = (\alpha, \beta) \in \Theta \subseteq \mathbb{R}^2$ $\theta_0 = (\alpha_0, \beta_0)$ is the true parameter value

High frequency asymptotics

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

$\theta = (\alpha, \beta) \in \Theta \subseteq \mathbb{R}^2$ $\theta_0 = (\alpha_0, \beta_0)$ is the true parameter value

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

High frequency asymptotics

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

$\theta = (\alpha, \beta) \in \Theta \subseteq \mathbb{R}^2$ $\qquad$ $\theta_0 = (\alpha_0, \beta_0)$ is the true parameter value

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

High frequency asymptotic scenario:

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

High frequency asymptotics: conditions

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

Condition for rate-optimality:

Jacobsen's condition:

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

for all x and $\theta \in \Theta$

High frequency asymptotics: conditions

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

Condition for rate-optimality:

Jacobsen's condition:

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

for all x and $\theta \in \Theta$

Condition for efficiency:

$$\partial_y g_1(0, x, x; \theta) = \partial_\alpha b(x; \alpha)/\sigma^2(x; \beta) \qquad \partial_y^2 g_2(0, x, x; \theta) = \partial_\beta \sigma^2(x; \beta)/\sigma^4(x; \beta)$$

for all x and $\theta \in \Theta$

High frequency asymptotics: conditions

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

Condition for rate-optimality:

Jacobsen's condition:

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

for all x and $\theta \in \Theta$

Condition for efficiency:

$$\partial_y g_1(0, x, x; \theta) = \partial_\alpha b(x; \alpha)/\sigma^2(x; \beta) \qquad \partial_y^2 g_2(0, x, x; \theta) = \partial_\beta \sigma^2(x; \beta)/\sigma^4(x; \beta)$$

for all x and $\theta \in \Theta$

Jacobsen (2001): small Δ-optimality

High frequency asymptotics

The asymptotic distribution of the optimal estimator is

$$\begin{pmatrix} \sqrt{n\Delta_n}(\hat{\alpha}_n - \alpha_0) \\ \\ \sqrt{n}(\hat{\beta}_n - \beta_0) \end{pmatrix} \xrightarrow{\mathcal{D}} N_2\left(0, \mathcal{J}^{-1}\right)$$

where

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

μ_θ is the density of the invariant measure

By the LAN-result in Gobet (2002), the estimator is efficient and rate-optimal

High frequency asymptotics

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

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

$a_j^*(x, \Delta; \theta)$ optimal weights

Rate-optimality and efficiency follows under regularity conditions including ergodicity of X, smoothness of the functions f_j, $N \geq 2$, and that the matrix

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

is invertible for all x