

Eigenfunctions based estimating martingales for perturbed diffusions

Martingale basate su autofunzioni per la stima di diffusionsi perturbate

Francesco Corielli

Istituto di Metodi Quantitativi

Università Bocconi, Milano

Francesco.Corielli@Uni-Bocconi.it

Riassunto: In questa nota si estende il metodo proposto in (Kessler and Sorensen 1999) per la stima di parametri in diffusionsi osservate a tempo discreto a una vasta classe di modelli perturbati. Come esempio e' illustrato un test per l'appartenenza di una diffusione ad una data classe polinomiale.

Keywords: Diffusion processes, Martingale functions.

1. Introduction

The estimation of parameters for diffusion processes is a relevant applied problem with implications for stochastic modelling in such different subjects as physics, economics, meteorology and hydrology. An elegant and classical estimation theory exists for the case of processes observable in continuous time. Many attempts have been made to extend the theory to the, more realistic, discrete time case. This is not an easy task as, if we want to avoid gross approximations, it requires the solution of complex problems expressed in terms of partial differential equations (PDE). The introduction of the concept of martingale estimating functions has allowed the research to move fast in the direction of effectively implementable algorithms (see e.g. (Godambe and Heyde 1987) and the review in (Sorensen 1997)). (Kessler and Sorensen 1999) (KS) pointed the way toward a systematic exploitation of classical Sturm Liouville (SL) theory (in particular in the polynomial case) for computing martingale functions based on eigenfunctions and eigenvectors of the diffusion generator. However, when eigenvectors and eigenfunctions are known, the transition density and, hence, the likelihood function are known too, at least as eigenexpansions. So, a more thorough extension of known estimation methods can be achieved only by considering problems where eigenfunction and eigenvalues are not explicitly known. In this note we exploit the completeness of polynomial solutions of the SL problem as a system of orthogonal functions in a suitable weighted space of functions (the weight being the invariant probability for the polynomial diffusion) and apply the technique of perturbative solution for PDEs, in order to derive easily computable approximate martingale functions for a wide spectrum of perturbed polynomial diffusions. As an application, we describe a test for the null hypothesis that an observed diffusion belongs into a given polynomial class.

2. Martingale functions based on eigenfunctions

In order to save space, in this short note we share the hypothesis set of (Kessler and Sorensen 1999) to which we refer for details. Let X_t be a solution for $0 \leq t \leq T$ of the (time homogeneous) SDE

$$dX = b_\theta(X)dt + \sigma_\theta(X)dW \quad X_0 = x_0 \quad (1)$$

Where θ is a vector of real valued parameters, W is a one dimensional Brownian motion and x_0 is a constant. We suppose that the solution of 1 exists at least in the weak sense. As in KS we suppose that b and σ are such that the diffusion X is ergodic and admits a unique invariant density μ . The solutions of the problem

$$g_x b + \frac{1}{2} g_{xx} \sigma^2 + g_t = 0 \quad (2)$$

in the class of square integrable ($\forall t$ w.r.t. μ) functions $g \equiv g_\theta(X, t)$ are martingale functions of the process X . If we suppose $g_\theta(X, t) = h_\theta(X)k_\theta(t)$ the differential problem becomes an eigenproblem of the Sturm Liouville (SL) type for the generator of the diffusion, that is

$$h_x b + \frac{1}{2} h_{xx} \sigma^2 + h\lambda = 0; \quad \lambda = k_t/k \quad (3)$$

where $\lambda \equiv \lambda_\theta$. Under the hypothesis of KS, this problem, with proper initial-boundary conditions (not necessarily homogeneous as in standard SL theory), has (martingale) solutions

$$e^{-\lambda_j t} h_j(X) = g_j(X, t) \quad (4)$$

where (avoiding, for simplicity of notation, to explicit the dependence on θ) $\lambda_1 < \lambda_2 < \dots$ is a denumerable set of eigenvalues and h_j are the corresponding eigenfunctions. If we are able to explicitly solve this problem, we can derive moment conditions for estimating θ . In fact, since

$$E_{t_i} [h_j(X_{t_{i+1}}) - e^{-\lambda_j(t_{i+1}-t_i)} h_j(X_{t_i})] = 0 \quad (5)$$

we can suggest, for a sample $x_{t_1}, x_{t_2}, \dots, x_{t_n}$ an estimate $\hat{\theta}$ based on the empirical equivalent of 5 for any set of r solutions of 3, that is

$$\hat{\theta}_n : \sum_{i=1}^{n-1} \sum_{j=1}^r [h_j(x_{t_{i+1}}) - e^{-\lambda_j(t_{i+1}-t_i)} h_j(x_{t_i})] = 0 \quad (6)$$

Using standard martingale results, KS show the consistency and asymptotic normality of such estimates. However, for KS methods to be applicable the eigenfunctions must be explicitly known (KS consider in detail the case of polynomial eigenfunctions). This is a limit of the analysis, since when the eigenfunctions and eigenvalues are known, transition densities have a series expansion in terms of these and, at least in principle, maximum likelihood estimates are available.

3. Perturbed polynomial diffusions

In order to extend KS results to a bigger set of models where eigenfunctions cannot be given explicitly, we consider polynomial solutions of the SL problem and recall the fact that these constitute complete orthogonal systems in the space of square integrable functions with respect to the corresponding μ . We can, then, look for an expansion of non polynomial solutions in terms of polynomial eigenfunctions. The problem here is that of deriving the proper expansion without actually computing the solution of the non polynomial problems. A useful tool for solving this problem is perturbation theory. For the sake of simplicity let us concentrate on the diffusion function σ and let us consider a polynomial problem, perturbed by distorting its diffusion in the direction of a non polynomial problem q . Notice that q shall, in general, depend on unknown parameters. In this case we add these parameters to the vector θ which, as in the previous section, for simplicity we drop from the notation. Let us write the perturbed diffusion as

$$dX^\varepsilon = b(X^\varepsilon)dt + \sqrt{\sigma^2(X^\varepsilon) + \varepsilon [q^2(X^\varepsilon) - \sigma^2(X^\varepsilon)]}dW \quad (7)$$

Where $0 \leq \varepsilon \leq 1$. If we suppose that X^ε exists for $\varepsilon = 0$ and $\varepsilon = 1$ then it shall exist for all $0 \leq \varepsilon \leq 1$ so that we can meaningfully speak of martingale solutions for 7. The corresponding SL problem shall be

$$h_x^\varepsilon b + \frac{1}{2} h_{xx}^\varepsilon [\sigma^2 + \varepsilon [q^2 - \sigma^2]] + h^\varepsilon \lambda^\varepsilon = 0; \quad \lambda^\varepsilon = k_t^\varepsilon / k^\varepsilon$$

If we suppose that the solution of this problem: $g^\varepsilon = h^\varepsilon k^\varepsilon$, can be expanded in powers of ε : $g^\varepsilon = g_0 + \varepsilon g_1 + \varepsilon^2 g_2 + \dots$ where g_0 is the solution of the unperturbed problem and set $P^2 = q^2 - \sigma^2$, we derive a chain of PDEs

$$\begin{aligned} g_{0x}b + 1/2\sigma^2 g_{0xx} + g_{0t} &= 0 \\ g_{1x}b + 1/2\sigma^2 g_{1xx} + g_{1t} &= -1/2P^2 g_{0xx} \\ g_{2x}b + 1/2\sigma^2 g_{2xx} + g_{2t} &= -1/2P^2 g_{1xx} \\ &\dots \end{aligned} \quad (8)$$

where the PDE for g_0 is to be solved under the initial/boundary conditions of the unperturbed problem, while the other PDEs must be solved under null initial conditions. The solution can be expressed in terms of eigenvalues λ_j^ε and eigenfunctions h_j^ε expansions:

$$\begin{aligned} \lambda_j^\varepsilon &= \lambda_j + \varepsilon \lambda_j^{(1)} + \varepsilon^2 \lambda_j^{(2)} + \dots ; h_j^\varepsilon = h_j + \varepsilon h_j^{(1)} + \varepsilon^2 h_j^{(2)} + \dots \\ \lambda_j^{(1)} &= (h_j, H h_j) \\ \lambda_j^{(2)} &= (h_j, [H - \lambda_j^{(1)}] h_j^{(1)}) \\ \lambda_j^{(3)} &= (h_j^{(1)}, [H - \lambda_j^{(1)}] h_j^{(1)}) - \lambda_j^{(2)} (h_j, h_j^{(1)} + (h_j^{(1)}, h_j^{(0)})) \\ &\dots \end{aligned} \quad (9)$$

$$h_j^{(1)} = \sum_{m \neq j} \frac{(h_m, Hh_j)}{\lambda_j - \lambda_m} h_m \quad (10)$$

$$h_j^{(2)} = \sum_{m \neq j} \sum_{l \neq j} \frac{(h_m, Hh_l)}{\lambda_j - \lambda_m} \frac{(h_l, Hh_n)}{\lambda_j - \lambda_l} h_m - (h_j, Hh_j) \sum_{m \neq i} \frac{(h_m, Hh_j)}{\lambda_j - \lambda_m} h_m$$

...

Having set $H = P^2 \frac{\partial^2}{\partial x^2}$ and $(a, Hb) = \int_X a(x) P^2 \frac{\partial^2}{\partial x^2} b(x) \mu(x) dx$. A truncation of the series for λ_j^ε and h_j^ε shall give us approximate martingale functions for estimating the parameter vector θ and the perturbation parameter ε . Theorems on the consistency and asymptotic normality of the estimates of θ and ε can be derived, in a similar way as in KS, on the basis of classical martingale convergence results. The estimate of ε shall be useful for assessing whether data are compatible with the unperturbed polynomial model or reject it for the perturbed version.

4. A test for polynomial models

The perturbation results of the above section allow us to derive a simple test against the null of a given polynomial model. In order to illustrate the test in minimal space, we suppose that the functions b, σ and q are fully known, avoiding the need for estimating nuisance parameters. The extension to the nuisance parameters case is straightforward, but cumbersome in notation. If we estimate ε using the approximate martingale functions given in the previous section, we can state the following, slight, extension of Theorem 4.3 in (Kessler and Sorensen 1999)

Proposition: Let $\hat{\varepsilon}_n$ be the estimate of ε derived from condition 6 where λ_j and h_j are replaced by ${}_\varepsilon \lambda_j$ and ${}_\varepsilon h_j$ truncated after terms of order one in ε . Then, under $H_0 : \varepsilon = 0$, $\sqrt{n} \hat{\varepsilon}_n$ converges in distribution to a gaussian R.V. with expected value 0 and variance v/f^2 with $v = \sum_{j,m=1}^r \int_X a_{j,m}(x) \mu(x) dx$ and $f = \sum_{j=1}^r \int_X \pi(y, x, \Delta) \mu(x) \gamma_j(y, x) dx dy$, where $a_{j,m}(x) = \int_Y h_j(y) h_m(y) \pi(y, x, \Delta) dy - e^{-(\lambda_j - \lambda_m)\Delta} h_j(x) h_m(x)$, $\pi(x, y, \Delta)$ is the (known) transition density of the polynomial model, Δ is the time (here assumed constant) between consecutive observations and $\gamma_j(y, x) = h_j^{(1)}(y) - e^{-\lambda_j^{(1)} \Delta} h_j^{(1)}(x)$.

Notice that the values of v and f are easily computed as linear functions of the (known) moments of the polynomial process. A proper $1 - \alpha$ critical region for $H_0 : \varepsilon = 0$ shall then be $C_T \equiv \{x_{t_1}, \dots, x_{t_n}\} : \hat{\varepsilon}_n \notin [\pm z_{1-\alpha/2} \sqrt{v/nf^2}]$.

References

- Godambe, V. P. and Heyde, C. C. (1987). Quasi likelihood and optimal estimation, *Int. Statist. Rev.* **55**, 231–244.
- Kessler, M. and Sorensen, M. (1999). Estimating equations based on eigenfunctions for a discretely observed diffusion process, *Bernoulli* **5**, 299–314.
- Sorensen, M. (1997). Estimating functions for discretely observed diffusions: A review, *W.P. Aarhus University* pp. 1–21.