Inference from diffusion driven models based on estimating functions

Julie Lyng Forman
Department of Biostatistics. University of Copenhagen. Denmark.

Keywords: general estimating function, goodness of fit testing, stochastic differential equation.

Abstract

Diffusion-type models provide a flexible framework for modeling phenomena that evolve continuously and randomly with time. However, their statistical analysis is complicated as the functional form of the likelihood function seldom is known. General estimating functions often present a simple alternative means for fitting these models. This paper illustrates the use of estimating functions for diffusion-type models driven by Pearson diffusions, i.e. diffusions having linear drift and quadratic squared diffusion coefficient. Further a goodness of fit test based on downsampled estimating equations is presented.

1. Introduction

In this paper we consider first inference from a stationary diffusion process

$$dX_t = b(X_t, \psi)dt + \sigma(X_t, \psi)dB_t \tag{1}$$

with unknown parameter ψ based on a discrete-time sample $\{X_{i\Delta}\}_{i=1,\dots,n}$. Next, we consider more complex models driven by diffusions. The functional form of the likelihood function for theses model is hardly ever known and evaluating it by means of numerical methods is quite involved, see [5]

2. Estimating functions

General estimating functions provide a simple means for obtaining fast estimators from diffusion-type models. An estimating function is a function of the parameter and the data, $F(\psi) = F(X_\Delta, \dots, X_{n\Delta}, \psi)$ from which an estimate is found by solving the related estimating equation $F(\psi) = 0$. Often the estimating function takes the form,

$$F(\psi) = \sum_{i=r}^{n} f(X_{(i-r+1)\Delta}, \dots X_{i\Delta}, \psi) \tag{2}$$

where $E_\psi\{f(X_\Delta, \dots X_{r\Delta}, \psi)\} = 0$ corresponding to fitting certain moments of the process. The estimator is equivalent to the *generalized method of moments* estimator found by minimizing the criterion $|F(\psi)|^2$.

For a large class of diffusions explicit estimating functions can be based on the eigenfunctions of the infinitesemmal generator of the diffusion, see [4]. Recall that ϕ is an eigenfunction of the generator of the diffusion (1) if

$$b(x,\psi)\phi'(x,\psi) + (1/2)\sigma^2(x,\psi)\phi''(x,\psi) = -\lambda(\psi)\phi(x,\psi)$$

where $-\lambda(\psi)$ is the corresponding eigenvalue. This implies that

$$E_\psi\{\phi(X_{(i+1)\Delta},\psi)|X_{i\Delta}\} = \exp\{-\lambda(\psi)\Delta\}\phi(X_{i\Delta},\psi). \quad (3)$$

The estimating function proposed by [4] takes the form

$$F(\psi) = \sum_{i=2}^{n}\sum_{j=1}^{k} w_j(X_{i\Delta},\psi)[\phi_j(X_{(i+1)\Delta},\psi) - \exp\{-\lambda_j(\psi)\Delta\}\phi_j(X_{i\Delta},\psi)] \quad (4)$$

where $\phi_1,\ldots,\phi_k$ are eigenfunctions and $w_1,\ldots,w_k$ are suitable weight matrices. The estimating function has the interpretation of an approximation to the score and thus often produce good estimates.

3. The Pearson diffusions

Some highly tractable examples are given by the class of Pearson diffusions, see [1] and [6]. These are the diffusions

$$dX_t = -\theta(X_t-\mu)dt + \sqrt{2\theta(aX_t^2+bX_t+c)}dB_t \quad (5)$$

having mean reverting linear drift (i.e. $\theta > 0$) and quadratic squared diffusion coefficient. In our paper [1] we have identified all of the stationary Pearson diffusion dividing them into six subclasses corresponding to the form of the squared diffusion coefficient. An overview is presented in table 1. The invariant

	$\sigma^2(x)$	State space	Restrictions	Invariant distr.
1	2θ	$\mathbb{R}$	none	Normal
2	$2\theta x$	$]0;\infty[$	$\mu > 0$	Gamma
3	$2\theta a(x^2+1)$	$\mathbb{R}$	$a > 0$	(skewed) T
4	$2\theta ax^2$	$]0;\infty[$	$a,\mu > 0$	inverse Gamma
5	$2\theta ax(x+1)$	$]0;\infty[$	$a,\mu > 0$	(scaled) F
6	$2\theta ax(x-1)$	$]0;1[$	$a < 0,\ 0 < \mu < 1$	Beta

Table 1: Standard-type stationary Pearson diffusions. As the Pearson subclasses are closed under translation and transformations of scale any other stationary Pearson diffusion can be linked to its standard-type by a linear transformation. See [1] for details.

distributions all belong to the Pearson system. Hence, the name the Pearson diffusions suits the class. Most of the Pearson diffusions such as the Ornstein-Uhlenbeck processes (class 1) and the square root processes (class 2) are well known, but the class 3 and class 5 diffusions are yet to be explored. The Pearson diffusions have polynomial eigenfunctions $p_m(x) = \sum_{j=0}^{n} p_{m,j} x^j$ with coefficients recursively defined by.

$$(a_j - a_m)p_{m,j} = b_{j+1}p_{m,j+1} + c_{j+2}p_{m,j+2} \tag{6}$$

where $a_j = j\{1-(j-1)a\}$, $b_j = j\{\mu + (j-1)b\}$, and $c_j = j(j-1)c$. Assuming $p_{m,m} = 1$ and $p_{m,m+1} = 0$, then the coefficients $\{p_{m,j}\}_{j=0,\ldots,m-1}$ form a unique solution to the linear system (6). The eigenvalues are given by $\lambda_m = a_m\theta$. Hence, the discretely observed diffusions are easily fitted by means of estimating functions of the form (4). It is important to notice that explicit expressions of the conditional moments of the Pearson diffusions can be derived from (3) when inserting the polynomial eigenfunctions,

$$E(X_t^m \mid X_0 = x) = e^{-a_m t} \sum_{j=0}^{m} p_{m,j} x^j - \sum_{j=0}^{m-1} p_{m,j} E(X_t^j | X_0 = x). \tag{7}$$

This in turn can be used to calculate moments and mixed moments.

4. Diffusion driven models

The Pearson diffusions are particularly useful as input in more complex continuous time models. Examples of such are summed diffusions, integrated diffusions and stochastic volatility models. When based on the Pearson diffusions these models can be fitted by explicit moment conditions.

The summed diffusions constitute a flexible class of stochastic process models which for instance fit turbulence data well. The joint moments of independent summed Pearson diffusions $X_t = X_{1,t} + \ldots + X_{m,t}$ satisfy

$$E(X_s^k X_t^\ell) = \sum\sum \binom{k}{k_1, \ldots, k_m}\binom{\ell}{\ell_1, \ldots, \ell_m} E(X_{1,s}^{k_1} X_{1,t}^{\ell_1}) \cdots E(X_{m,s}^{k_m} X_{m,t}^{\ell_m})$$

where the summation is over $k_1, \ldots, k_m \geq 0$ such that $k_1 + \ldots + k_m = k$ and similarly for the ℓ's. Due to (7) these expressions can be made explicit.

Integrated observations,

$$Y_i = \frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} X_s \, ds,$$

occur when a diffusion can only be observed averaged over certain time periods. The joint moments of the integrated diffusion are determined by

$$E(Y_i^k Y_j^\ell) = \Delta^{-(k+\ell)} \int_A E\{X_{s_1} \cdots X_{s_k} X_{t_1} \cdots X_{t_\ell}\} ds_1 \ldots ds_k dt_1 \ldots dt_\ell.$$

where the domain of integration $A = [(i-1)\Delta, i\Delta]^k \times [(j-1)\Delta, j\Delta]^\ell$ can be reduced considerably by symmetry arguments. In case of an integrated Pearson diffusion an explicit expression is found when applying (7) repeatedly.
A stochastic volatility model is a generalization of the Black-Scholes model for the logarithm of an asset price

$$dX_t = (\kappa + \beta v_t)dt + \sqrt{v_t}dW_t$$

that takes into account the empirical finding that the variance varies randomly over time by for instance modeling the unobservable volatility process $\{v_t\}_{t\geq 0}$ by a positive diffusion. The observed returns $Y_i = X_{i\Delta} - X_{(i-1)\Delta}$ are conditionally independent and normally distributed given the volatility, with mean $M_i = \kappa\Delta + \beta S_i$ and variance $S_i = \int_{(i-1)\Delta}^{i\Delta} v_t dt$. If $\{v_t\}_{t\geq 0}$ is modeled as a Pearson diffusion or a sum of Pearson diffusions, then the joint moments can be explicitly calculated as for instance $E(Y_i^k Y_j^\ell)$ equals

$$\sum\sum \binom{k}{k_1\ k_2\ k_3}\binom{\ell}{\ell_1\ \ell_2\ \ell_3}(\kappa\Delta)^{k_1+\ell_1}\beta^{k_2+\ell_2}\zeta_{k_3}\zeta_{\ell_3}E(S_i^{k_2+k_3/2}S_j^{\ell_2+\ell_3/2}),$$

where the sum is over integers $k_1, k_2, k_3 \geq 0$ such that $k_1 + k_2 + k_3 = k$ and similarly for the ℓ's. The constant ζ_m is the m'th order moment of the standard normal distribution. Note that $\zeta_m = 0$ when m is odd. Hence, the problems reduces to finding the joint moments of an integrated diffusion as in the above.

5. Goodness of fit testing

When considering moment based inference it is natural to check excess moment conditions by use of the GMM overidentifying restrictions test, see [3]. In [2] we have investigated a GMM-type test aimed at checking the dependence structure in diffusion driven models. For this we assume that the model is fitted by an estimating function of the form (2) satisfying $Ef(X_\Delta, \ldots, X_{r\Delta}, \psi) = 0$ if and only if $\psi = \psi_0$.
The idea for the goodness of fit test is to check the consistency of the estimates when downsampling the data. To this end we consider the *downsampled* estimating functions F_k given by

$$F_k(\psi) = \sum_{i=1}^{n-rk} f(X_{i\Delta}, X_{k+i\Delta}, \ldots, X_{rk+i\Delta}, \psi).$$

Note that F_k is based on all observations, not only every k'th. I.e. F_k is the sum of k estimating functions based on different downsamples. This seems beneficial as no observations are wasted. Denote by $\hat{\psi}_k$ the estimator associated to F_k. In case the model is well specified, $\hat{\psi}_1$ and $\hat{\psi}_k$ both are consistent estimators of ψ_0. We would thus expect that $\hat{\psi}_1 \approx \hat{\psi}_k$ for n sufficiently large. Hence, we would reject the hypothesis when observing a large value of

$$\tau(k) = n^{-1} \cdot (\hat{\psi}_1 - \hat{\psi}_k) W_n^{-1} (\hat{\psi}_1 - \hat{\psi}_k)^T. \tag{8}$$

Under standard regularity conditions $\hat{\psi}_1 - \hat{\psi}_k$ is asymptotically normal and W_n should be an estimate of the asymptotic covariance matrix implying that $\tau(k) \xrightarrow{\mathcal{D}} \chi_d^2$ if the model is well specified. To be specific the asymptotic covariance is given by $\Sigma_0 = S_0^T \Gamma_0 S_0$, where $\Gamma_0 = \lim_{n\to\infty} n^{-1} \operatorname{Var}(\{F_1(\psi_0), F_k(\psi_0)\})$ and $S_0 = E(\partial_\psi \{F_1(\psi_0), F_k(\psi_0)\})$, which can be estimated by a sandwich covariance estimator. All in all the test is easy to compute as it only requires the evaluation of the estimating function used for fitting the model.
Our numerical examples suggest that the test is successful in distinguishing plain diffusions from non-markovian diffusion-type models, but the power can also be quite poor when comparing different diffusions.

Acknowledgements: The research is joint work with Michael Sørensen, Department of Mathematical Sciences, University of Copenhagen and with Bo Markusen and Helle Sørensen, Department of Natural Sciences, LIFE, University of Copenhagen.

6. Bibliography

[1] Forman, J. L. and Sørensen, M. (2006) *The Pearson Diffusions: A class of Statistically Tractable Diffusion Processes.* Preprint no 7, Department of Mathematical Science, University of Copenhagen.

[2] Forman, J. L. and Markusen, B. and Sørensen, H. (2007) *Goodness of fit based on downsampling with applications to continuous-time processes.* Working paper, Department of Mathematical Science, University of Copenhagen.

[3] Hall, A. R. (2005) *Generalized Method of Moments.* Oxford University Press.

[4] Kessler, M. and Sørensen, M. (1999) *Estimating equations based on eigenfunctions for a discretely observed diffusion.* Bernoulli, vol. 5, 299–314.

[5] Pitt, M. and Chib, S. and Shephard, N. (2006) *Likelihood based inference for diffusion driven state space models.* Working paper, Nuffield College, Oxford.

[6] Wong, E. (1964) *The Construction of a Class of Stationary Markoff Processes.* In Bellman, R., editor, *Stochastic Processes in Mathematical Physics and Engineering,* 264–276. American Mathematical Society, Rhode Island.