APPROXIMATION OF CONDITIONAL MOMENTS OF DIFFUSION PROCESSES

Isao Shoji
Institute of Policy and Planning Sciences, University of Tsukuba
Tsukuba Ibaraki 305-8573, Japan
e-mail:shoji@sk.tsukuba.ac.jp

ABSTRACT. A computational method of conditional moments of d–dimensional diffusion processes is considered. We show that conditional moments of a diffusion process can be approximated by the solution to an ordinary differential equation whose coefficients are characterized by coefficients of the stochastic differential equation of the process. We also show that the method gives the exact conditional moments if a stochastic differential equation has a special form. The numerical experiment of estimating parameters of a nonlinear stochastic differential equation from discrete observation shows that the maximum likelihood estimation with the computed conditional moments performs better than a conventional discretization approach.

AMS Subj. Classification: 65C50; 65C60
Key Words: Approximation; Conditional moments; Diffusion process; Ordinary differential equation; Stochastic differential equation.

1. Introduction

Stochastic behaviors of a diffusion process are characterized by its transition distribution function induced by the stochastic differential equation of the process. Nevertheless some statistics, say, the conditional mean and variance, are often used to describe the behaviors. In practice the conditional mean is used to predict the future value of the process and the conditional variance is used to its diversity. Recently in finance knowledge of the conditional variance particularly plays an important role in modeling financial time series. The conditional mean and variance, representatives of conditional moments, give important information on dynamics of the process. More directly conditional moments are used to derive a pricing model of a derivative security because the price is formulated as the expectation of a pay off at maturity. From a statistical point of view, conditional moments are also useful for estimating parameters of a diffusion process. For example, in the maximum likelihood estimation or quasi-maximum likelihood estimation, the conditional mean and variance are necessary to set up a likelihood function. In the generalized method of moment estimation proposed by Hansen [6], conditional moments are used to construct orthogonality conditions which are crucial for the method.

Date: 28 May, 2001.

In this way conditional moments of diffusion processes are widely used for many practical purposes. However, they are not easily computed.

One of most straightforward ways for the computation is to use the transition probability density function (PDF). PDF is given as a solution to the well-known Fokker-Planck equation. Conditional moments are given by integration under the PDF. This, however, is not a practical way. Except for very simple diffusion processes a closed form solution to the Fokker-Planck equation can not be obtained. For example, Wong [14] shows that even one-dimensional case, some restrictions on coefficients of the Fokker-Planck equation are required for an explicit formula. Thus some numerical methods are needed in general. These methods are often much time consuming and not easily to manipulate, particularly in multi-dimensional case. Formidably, numerical integration under the PDF given numerically is further needed to obtain conditional moments.

Another way is discretization of a stochastic differential equation (SDE) of a diffusion process. For example, applying the Euler scheme which assumes coefficients of a SDE to be piecewisely constant to the SDE, an approximate discretized process of the diffusion process is easily obtained. More conveniently, since the discretized process is expressed as sum of a measurable function at current time and an innovative process distributed normally, its conditional mean and variance are derived from the discrete process. However, since it is well-known that approximation by the Euler scheme is not efficient unless discrete time interval is very short, the approximate conditional mean and variance are probably far from the true ones. Of course, a more efficient approximation scheme may give the more accurate conditional mean and variance. However, the scheme almost always leads to a specific form of discretization; for example, see Kloeden and Platen [8], Biscay et al. [3] and Shoji [11]. Thus conditional moments, especially higher order, may not be easily computed. Furthermore, since evaluation of how much conditional moments are close to the true ones is different from that of how much sample paths of an approximate process are close to the true, a theoretical study about the former evaluation is inevitable. We must say the discretization approach is slow to compute conditional moments. Thus a more direct method is worth investigating.

The aim of the paper is to present a computational method of conditional moments of d–dimensional diffusion processes and to study theoretical aspects of the method. In general the proposed method is to compute conditional moments approximately. However, the method gives exact conditional moments if a SDE has a special form. As one of numerical examples an application to estimation problem is considered. From practical viewpoints, estimating parameters of a diffusion process from discrete observation is very important since real data are almost always observed at discrete times. There are many studies about this subject; for example, Yoshida [15], Bibby and Sorensen [2], Pedersen [9], and Prakasa Rao [10]. By using the conditional mean and variance given by the proposed method, a quasi-maximum likelihood estimation is carried out. We compare the estimation performance with a conventional estimation.

The paper is organized as follows. In section 2 we present a computational method of diffusion processes with theoretical results which give accuracy of approximation. In section 3 two numerical experiments are presented; one is a comparison of computed conditional moments by the proposed method and a conventional one using a SDE with linear drift and state dependent diffusion coefficient. The other is a comparison of performance of estimation for a nonlinear SDE. The technical details are provided in the appendix.

2. Computational method of conditional moments

Let $(\Omega, P, \mathcal{F})$ be a probability space and $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration. On the space, we consider a continuous adapted d–dimensional diffusion process X_t which satisfies the following stochastic differential equation:

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

where, $f \in C^\infty(R^d, R^d)$, $\sigma \in C^\infty(R^d, R^d \times R^r)$, and B_t is an r–dimensional standard Brownian motion. In this paper X_t is considered on $t \in [s, \infty)$ and so B_t is a Brownian motion on $\{\mathcal{F}_t\}_{t\geq s}$ starting at B_s.

In practice we are only interested in bounded processes since unbounded ones are out of control. However, X_t defined above is not always bounded. In order to restrict the process to be bounded, we introduce the following stopping time,

$$T = \inf\{t \geq s; \|X_t - X_s\| \geq K_X\}, \tag{2}$$

where K_X is a finite positive constant. Let $\tilde{X}_t = X_{t\wedge T}$. Clearly, $\tilde{X}_t$ is bounded and satisfies the following:

$$\tilde{X}_t = \begin{cases} X_s + \int_s^t f(X_u)du + \int_s^t \sigma(X_u)dB_u & \text{if } s \leq t \leq T, \\ X_T & \text{if } t > T. \end{cases}$$

From now on, we consider conditional moments of $\tilde{X}_t$ instead of X_t itself. Moreover, since we are interested in stochastic behaviors of the process for relatively short time interval, we will study conditional moments of the process at time $(s+\Delta t)$ where $\Delta t < 1$. The choice of time interval may affect possibility of events related to T. To characterize the relationship between T and Δt, we introduce a positive constant δ_X which basically represents a small number,

$$\delta_X = P\{T < s + \Delta t\}. \tag{3}$$

In other words, we make K_X large enough for δ_X to be small.

Under these settings, we define n–th conditional moment of $\tilde{X}_t$, $\psi_\alpha(t)$, with a multiple index $\alpha = (\alpha_1, \cdots, \alpha_d)$ of length $n = |\alpha| = \alpha_1 + \cdots + \alpha_d$,

$$\psi_\alpha(t) = E\left[\prod_{i=1}^{d}(\tilde{X}_{i,t} - \tilde{X}_{i,s})^{\alpha_i}\middle|\mathcal{F}_s\right], \tag{4}$$

where $\tilde{X}_{i,t}$ denotes i–th element of $\tilde{X}_t$ and of course $\tilde{X}_s = X_s$. Here we consider the moments around $\tilde{X}_s$ for simplicity in the computation. Now, suppose that we compute up to n–th conditional moments of $\tilde{X}_t$ $(0 \le t - s \le \Delta t)$.

Theorem 1. *Let $\Psi(t) = (\psi_\alpha(t))_{1\le|\alpha|\le n}$ which has dimension of $\phi(n)$. There exists an $\mathcal{F}_s$-measurable $(\phi(n) \times \phi(n))$ matrix A and $\mathcal{F}_s$–measurable $\phi(n)$–dimensional vector functions $b(t)$, $R(t)$, and $R_X(t)$ such that,*

$$\Psi(t) = A\int_s^t \Psi(u)du + b(t) + R(t) + R_X(t), \tag{5}$$

where $R(t)$ has order of $O((\Delta t)^{(n+3)/2})$ and $R_X(t)$ has $O(\Delta t)$ in Δt and $O(\delta_X)$ in δ_X.

Proof. The proof is given in the appendix.

In the theorem $R(t)$ means an approximation error associated with expansion of f and $\sigma\sigma^T$, and $R_X(t)$ does one associated with the stopping time T.

Theorem 2. *$\Psi(t)$ is give as,*

$$\Psi(t) = A\int_s^t \exp(A(t-u))b(u)du + b(t) + R'(t) + R'_X(t), \tag{6}$$

where $R'(t)$ and $R'_X(t)$ have the same orders as $R(t)$ and $R_X(t)$ respectively.

Proof. The proof is given in the appendix.

Theorem 2 states that conditional moments can be approximated by a solution to an ordinary differential equation whose coefficients are all $\mathcal{F}_s$-measurable:

$$\Psi(t) = A\int_s^t \Psi(u)du + b(t).$$

More specifically, the proof will show that A and $b(t)$ are characterized by partial derivatives of f and $\sigma\sigma^T$. Once the method is turned out to be an approximation one, we may ask whether or not the method can give exact conditional moments. This is true if (1) has a special form. In addition, there is no need to restrict X_t to be bounded by the stopping time. Instead let $\psi_\alpha(t) = E\left[\prod_{i=1}^{d}(X_{i,t} - X_{i,s})^{\alpha_i}\middle|\mathcal{F}_s\right]$.

Corollary 1. *If $f(X)$ is linear in X and $\sigma\sigma^T(X)$ is at most quadratic, then,*

$$\Psi(t) = A\int_s^t \exp(A(t-u))b(u)du + b(t). \tag{7}$$

Proof. The proof is given in the appendix.

It is interesting to see the similar relationship between the assumption of corollary 1 and the condition by Wong [14] that the transition probability density function has an expansion by orthogonal polynomials. Corollary 1 is also useful in the financial time series analysis since many financial models satisfies the assumption of the corollary; for example, see Cox, et al. [5] and Chan, et al. [4].

3. Numerical Experiments

We present two numerical examples of one–dimensional SDEs. The first SDE is often used for a financial model:

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dB_t.$$

Since the SDE satisfies the assumption of corollary 1, the proposed method gives the exact conditional moments. Here we compute the conditional mean and variance. It can be easily seen that,

$$\begin{aligned} A &= \begin{pmatrix} -\kappa & 0 \\ 2\kappa(\theta - X_s) + \sigma^2 & -2\kappa \end{pmatrix} \\ b(t) &= \begin{pmatrix} \kappa(\theta - X_s)(t-s) \\ \sigma^2 X_s(t-s) \end{pmatrix}. \end{aligned}$$

Then, the first and second moments of X_t are formulated,

$$\Psi(t) = A\int_s^t \exp(A(t-u))b(u)du + b(t).$$

Although eigen values of A are easily obtained because of its special form, we try to compute the integration numerically for general cases. For simplicity, the integration is replaced with the following finite summation,

$$\int_s^t \exp(A(t-u))b(u)du \approx \sum_{k=1}^{n} \exp(A(t-u_{k-1}))b(u_{k-1})(u_k - u_{k-1}),$$

where $s = u_0 < \cdots < u_n = t$ and $u_k - u_{k-1} = \Delta u$. Naturally this causes approximation error. In the following simulation we set $n = 8$. As for the summation, matrix exponentials can be computed by the formula in Bernstein and So [1].

To see the difference between the above method and a conventional one, discretization by the Euler scheme is considered; that is,

$$\begin{aligned} X_t &= X_s + \kappa(\theta - X_s)(t-s) + \varepsilon_t \\ \varepsilon_t &\sim N(0, \sigma^2 X_s(t-s)). \end{aligned}$$

This approximation is often used for financial modeling; for example, see Chan et al. [4].

On the other hand, from Cox, et al. [5] or Bibby and Sorensen [2], the exact conditional mean and variance of X_t are as follows:

$$\begin{aligned} E[X_t|\mathcal{F}_s] &= X_s\exp(-\kappa(t-s)) + \theta(1-\exp(-\kappa(t-s)) \\ E\left[(X_t - E[X_t|\mathcal{F}_s])^2\middle|\,\mathcal{F}_s\right] &= X_s\frac{\sigma^2}{\kappa}\left(\exp(-\kappa(t-s)) - \exp(-2\kappa(t-s))\right) \\ &\quad +\theta\frac{\sigma^2}{2\kappa}\left(1-\exp(-\kappa(t-s)\right)^2. \end{aligned}$$

Using the above formulation, we compare approximation error of the two methods through numerical simulation with $(\kappa, \theta, \sigma) = (0.5, 0.05, 0.1)$. Results of the simulation are displayed in figure 1. The solid line represents approximation error of the proposed method and the dashed one the discretization by the Euler scheme. Conditional moments were computed at every discrete time with $t-s = 0.1$. Sample points were generated with time interval $(t-s)/20$ by a local linearization method developed by Shoji [11]. The results show that the discrtization approach caused much lager error. On the other hand, an error of the proposed method can be reduced to zero theoretically as increasing n associated with the finite summation.

As the second we consider the following nonlinear SDE,

$$dX_t = aX_t^3dt + \sigma dB_t.$$

Here we want to estimate a and σ from discrete observation $\{X_{t_k}\}_{k=0}^N$. Since the likelihood function of X_t for discrete observation is not known, some approximation is needed. We will estimate the parameters by the quasi-maximum likelihood estimation (QML). Even though QML is not an exact one, some favorable properties are known. Their details are discussed in White [12, 13].

Let $t_k - t_{k-1} = \Delta t$. We use the following as a quasi-likelihood function,

$$L_Q(a,\sigma) = p(X_{t_0})\prod_{k=0}^{N-1}\frac{1}{\sqrt{2\pi V_{t_{k+1}|t_k}}}\exp\left(\frac{(X_{t_{k+1}} - X_{t_{k+1}|t_k})^2}{2V_{t_{k+1}|t_k}}\right),$$

where $X_{t_{k+1}|t_k}$ and $V_{t_{k+1}|t_k}$ are the conditional mean and variance which can be computed by the proposed method. It can be easily seen,

$$\begin{aligned} A &= \begin{pmatrix} 3aX_s^2 & 3aX_s \\ 2aX_s^3 & 6aX_s^2 \end{pmatrix} \\ b(t) &= \begin{pmatrix} aX_s^3(t-s) \\ \sigma^2(t-s) \end{pmatrix}. \end{aligned}$$

As an alternative, the discretization by the Euler scheme is used again. The likelihood function is as follows,

$$L_E(a,\sigma) = p(X_{t_0})\prod_{k=0}^{N-1}\frac{1}{\sqrt{2\sigma^2\Delta t}}\exp\left(\frac{\left(X_{t_{k+1}} - (X_{t_k} + aX_{t_k}^3\Delta t)\right)^2}{2\sigma^2\Delta t}\right).$$

We estimated a and σ with the true $(a, \sigma) = (-1, 1)$ and $\Delta t = 0.1$ for five hundred different paths each of which had one thousand discrete observations and formed empirical distribution of the estimates. In the same as the first example, sample points were generated and numerical integration was carried out. The empirical distributions are displayed in figure 2. Focusing on bias of the estimates, the QML approach shows smaller bias whereas the discretization approach overestimates a and considerably underestimates σ.

4. Concluding Remarks

In this paper, we have presented a computational method of conditional moments of d–dimensional diffusion processes and studied its theoretical aspects. It was shown that conditional moments can be approximated by a solution to an ordinary differential equation and also that the method gives the exact conditional moments if a stochastic differential equation has linear coefficients of drift and at most quadratic coefficients of quadratic variation. A numerical experiment was carried out to estimate parameters of a stochastic differential equation from discrete observation by the quasi-maximum likelihood estimation. The results of the experiments indicated that the proposed method shows smaller bias of estimates than a discretization approach. The method is also useful for other estimation methods, say, Hansen [6] or Bibby and Sorensen [2] which require conditional moments to construct object functions for estimation.

Appendix: Proofs for Section 2

In the following we assume $0 \le t - s \le \Delta t < 1$.

Lemma 1. *Let α be a multiple index of length n. For any integer n there exists a constant C_n, which depends only on n, such that,*

$$E\left[\prod_i (\tilde{X}_{i,t} - \tilde{X}_{i,s})^{2\alpha_i} \middle| \mathcal{F}_s \right] \le C_n (t-s)^n,$$

where for some constant K,

$$C_n < (2K^2 r)^n (n+2)!.$$

Proof. We prove the lemma by induction on n. We can assume $|\tilde{X}_{i,t}|$, $|f_i(X_t)|$, and $|\sigma_{i,j}(X_t)|$ are less than a constant K $(K \ge K_X)$ on $t \in [s, t \wedge T]$ because of $f_i \in C^\infty(R^d)$ and $\sigma_{i,j} \in C^\infty(R^d)$. Let $n = 1$. For each i Ito's formula gives,

$$
\begin{aligned}
(\tilde{X}_{i,t}-\tilde{X}_{i,s})^2 &= (X_{i,t\wedge T}-X_{i,s})^2 \\
&= \int_s^{t\wedge T} 2(X_{i,u}-X_{i,s})f_i(X_u)du + \sum_j \int_s^{t\wedge T} 2(X_{i,u}-X_{i,s})\sigma_{i,j}(X_u)dB_{j,u} \\
&\quad + \sum_j \int_s^{t\wedge T} \sigma_{i,j}^2(X_u)du.
\end{aligned}
$$

Applying the optional sampling theorem to the above,

$$
\begin{aligned}
& E\left[(X_{i,t\wedge T}-X_{i,s})^2|\mathcal{F}_s\right] \\
&\leq E\left[\int_s^{t\wedge T} 2|(X_{i,u}-X_{i,s})f_i(X_u)|du + \sum_j \int_s^{t\wedge T} \sigma_{i,j}^2(X_u)du \middle| \mathcal{F}_s\right] \\
&\leq E\left[\int_s^{t\wedge T} (X_{i,u}-X_{i,s})^2 du + \int_s^{t\wedge T} f_i^2(X_u)du + \sum_j \int_s^{t\wedge T} \sigma_{i,j}^2(X_u)du \middle| \mathcal{F}_s\right] \\
&\leq E\left[\int_s^{t\wedge T} (X_{i,u}-X_{i,s})^2 du + \int_s^{t\wedge T} K^2 du + \sum_j \int_s^{t\wedge T} K^2 du \middle| \mathcal{F}_s\right] \\
&\leq E\left[\int_s^{t} (X_{i,u\wedge T}-X_{i,s})^2 du \middle| \mathcal{F}_s\right] + (1+r)\int_s^t K^2 du.
\end{aligned}
$$

Using the Gronwall's inequality, for example see Karatzas and Shreve [7],

$$
\begin{aligned}
E[(X_{i,t\wedge T}-X_{i,s})^2|\mathcal{F}_s] &\leq (1+r)K^2(t-s) + (1+r)K^2 \int_s^t (u-s)\exp(t-u)du \\
&= (1+r)K^2(t-s)\frac{\exp(t-s)-1}{t-s}.
\end{aligned}
$$

Since $t-s<1$, we can set $C_1 = 2(1+r)K^2$. This shows that the claim holds for $n=1$. Let $|\alpha| = n$ $(n>1)$. Now suppose that the claim holds for the case which is less than n. Again, using Ito's formula and the same argument as above, we get,

$$
\begin{aligned}
&E\left[\prod_i (X_{i,t\wedge T}-X_{i,s})^{2\alpha_i}\middle|\mathcal{F}_s\right]\\
&= E\left[\sum_i \int_s^{t\wedge T} 2\alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{2\alpha_j} f_i(X_u)du\middle|\mathcal{F}_s\right]\\
&\quad +\frac{1}{2}E\left[\sum_{i,k}{}^2 \int_s^{t\wedge T} 2\alpha_i(2\alpha_i-1)(X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{l\neq i}(X_{l,u}-X_{l,s})^{2\alpha_l}\sigma_{i,k}^2(X_u)du\middle|\mathcal{F}_s\right]\\
&\quad +\frac{1}{2}E\Bigg[\sum_{i\neq j,k}{}^3 \int_s^{t\wedge T} 2\alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-1}2\alpha_j(X_{j,u}-X_{j,s})^{2\alpha_j-1}\prod_{l\neq i,j}(X_{l,u}-X_{l,s})^{2\alpha_l}\\
&\qquad\qquad \sigma_{i,k}(X_u)\sigma_{j,k}(X_u)du\Bigg|\mathcal{F}_s\Bigg]\\
&\leq E\left[\sum_i \int_s^{t\wedge T} 2\alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{2\alpha_j}|(X_{i,u}-X_{i,s})f_i(X_u)|du\middle|\mathcal{F}_s\right]\\
&\quad +\frac{1}{2}E\left[\sum_{i,k}{}^2 \int_s^{t\wedge T} 2\alpha_i(2\alpha_i-1)(X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{l\neq i}(X_{l,u}-X_{l,s})^{2\alpha_l}\sigma_{i,k}^2(X_u)du\middle|\mathcal{F}_s\right]\\
&\quad +\frac{1}{2}E\Bigg[\sum_{i\neq j,k}{}^3 \int_s^{t\wedge T} 2\alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-1}2\alpha_j(X_{j,u}-X_{j,s})^{2\alpha_j-1}\prod_{l\neq i,j}(X_{l,u}-X_{l,s})^{2\alpha_l}\\
&\qquad\qquad |(X_{i,u}-X_{i,s})(X_{j,u}-X_{j,s})\sigma_{i,k}(X_u)\sigma_{j,k}(X_u)|du\Bigg|\mathcal{F}_s\Bigg]\\
&\leq 4K^2E\left[\sum_i \int_s^{t\wedge T} \alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{2\alpha_j}du\middle|\mathcal{F}_s\right]\\
&\quad +K^2E\left[\sum_{i,k}{}^2 \int_s^{t\wedge T} \alpha_i(2\alpha_i-1)(X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{l\neq i}(X_{l,u}-X_{l,s})^{2\alpha_l}du\middle|\mathcal{F}_s\right]\\
&\quad +K^2E\Bigg[\sum_{i\neq j,k}{}^3 \int_s^{t\wedge T} \alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-2}\alpha_j(X_{j,u}-X_{j,s})^{2\alpha_j-2}\prod_{l\neq i}(X_{l,u}-X_{l,s})^{2\alpha_l}\\
&\qquad\qquad ((X_{i,u}-X_{i,s})^2+(X_{j,u}-X_{j,s})^2)du\Bigg|\mathcal{F}_s\Bigg],
\end{aligned}
$$

where $\sum^2$ and $\sum^3$ mean the double and triple summations subject to the conditions implied by logical formula of their subscripts. In the above, we assume that every integrand with negative index is zero. For the first term in the last inequality, use the supposition of induction on n.

$$
\begin{aligned}
&E\left[\sum_i \int_s^{t\wedge T} \alpha_i (X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{2\alpha_j}du \,\middle|\, \mathcal{F}_s\right]\\
&\leq \sum_i \int_s^t \alpha_i C_{n-1}(u-s)^{n-1}du\\
&= C_{n-1}(t-s)^n.
\end{aligned}
$$

For the second term,

$$
\begin{aligned}
&E\left[\sum_{i,k}{}^2 \int_s^{t\wedge T} \alpha_i(2\alpha_i-1)(X_{i,u}-X_{i,s})^{2\alpha_i-2}\prod_{l\neq i}(X_{l,u}-X_{l,s})^{2\alpha_l}du \,\middle|\, \mathcal{F}_s\right]\\
&\leq \sum_{i,k}{}^2 \int_s^t \alpha_i(2\alpha_i-1)C_{n-1}(u-s)^{n-1}du\\
&\leq r(2\sum_i \alpha_i^2 - n)C_{n-1}\frac{(t-s)^n}{n}.
\end{aligned}
$$

For the third term,

$$
\begin{aligned}
&E\Bigg[\sum_{i\neq j,k}{}^3 \int_s^{t\wedge T} \alpha_i(X_{i,u}-X_{i,s})^{2\alpha_i-2}\alpha_j(X_{j,u}-X_{j,s})^{2\alpha_j-2}\prod_{l\neq i,j}(X_{l,u}-X_{l,s})^{2\alpha_l}\\
&\qquad \left((X_{i,u}-X_{i,s})^2+(X_{j,u}-X_{j,s})^2\right)du \,\Bigg|\, \mathcal{F}_s\Bigg]\\
&\leq 2r\sum_{i\neq j}{}^2 \int_s^t \alpha_i\alpha_j C_{n-1}(u-s)^{n-1}du\\
&= 2r(n^2-\sum_i \alpha_i^2)C_{n-1}\frac{(t-s)^n}{n}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
E\left[\prod_i(\tilde{X}_{i,t}-\tilde{X}_{i,s})^{2\alpha_i}\,\middle|\,\mathcal{F}_s\right] &\leq K^2(2nr-r+4)C_{n-1}(t-s)^n\\
&\leq 2K^2r(n+2)C_{n-1}(t-s)^n.
\end{aligned}
$$

The last inequality follows from $r\geq 1$. Define $C_n = 2K^2r(n+2)C_{n-1}$. Then,

$$
\begin{aligned}
C_n &= \frac{1}{6}C_1(2K^2r)^{n-1}(n+2)!\\
&= \frac{1}{6}\left(1+\frac{1}{r}\right)(2K^2r)^n(n+2)!\\
&< (2K^2r)^n(n+2)!.
\end{aligned}
$$

This completes the proof.

Proof of the Theorem 1. For simplicity $t-s=\Delta t$ in the following. Let $n\geq 1$ be fixed and $1\leq m\leq n$. Let α be a multiple index of length m. Applying the Ito's formula,

$$
\begin{aligned}
&E\left[\prod_i (X_{i,t\wedge T}-X_{i,s})^{\alpha_i}\middle|\mathcal{F}_s\right]\\
=\ & E\left[\sum_i\int_s^{t\wedge T}\alpha_i(X_{i,u}-X_{i,s})^{\alpha_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j}f_i(X_u)du\middle|\mathcal{F}_s\right]\\
&+\frac{1}{2}E\left[\sum_{i,k}{}^2\int_s^{t\wedge T}\alpha_i(\alpha_i-1)(X_{i,u}-X_{i,s})^{\alpha_i-2}\prod_{l\neq i}(X_{l,u}-X_{l,s})^{\alpha_l}\sigma_{i,k}^2(X_u)du\middle|\mathcal{F}_s\right]\\
&+\frac{1}{2}E\Bigg[\sum_{i\neq j,k}{}^3\int_s^{t\wedge T}\alpha_i(X_{i,u}-X_{i,s})^{\alpha_i-1}\alpha_j(X_{j,u}-X_{j,s})^{\alpha_j-1}\prod_{l\neq i,l\neq j}(X_{l,u}-X_{l,s})^{\alpha_l}\\
&\qquad\qquad \sigma_{i,k}(X_u)\sigma_{j,k}(X_u)du\Bigg|\mathcal{F}_s\Bigg].
\end{aligned}
$$

Now we consider separately the three parts of the right hand side. Applying the Taylor's expansion of $f_i(X)$ up to $(n-m+2)$–th order to the first part with multiple index β,

$$
\begin{aligned}
&E\left[\int_s^{t\wedge T}(X_{i,u}-X_{i,s})^{\alpha_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j}f_i(X_u)du\middle|\mathcal{F}_s\right]\\
=\ &\sum_{0\leq|\beta|\leq n-m+1}\frac{\partial^\beta f_i(X_s)}{|\beta|!}E\left[\int_s^{t\wedge T}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\middle|\mathcal{F}_s\right]\\
&+\sum_{|\beta|=n-m+2}E\left[\int_s^{t\wedge T}\frac{\partial^\beta f_i(\xi_u)}{|\beta|!}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\middle|\mathcal{F}_s\right].
\end{aligned}
$$

where $\partial^\beta=\partial^{|\beta|}/(\partial x_1)^{\beta_1}\cdots(\partial x_d)^{\beta_d}$ and $\xi_u=X_s+\theta_u(X_u-X_s)$ with $0\leq\theta_u\leq 1$. Paying our attention to the case $|\alpha|=m=1$, the conditional expectation in the main part of the Taylor's expansion is expressed as follows:

(1) $\alpha=e_i$ $(1\leq i\leq d)$ and $|\beta|=0$ where the i–th element of e_i is one and the others zeros: Then, $|\alpha+\beta-e_i|=0$ and so the integrand is one. The conditional expectation is $E[t\wedge T-s|\mathcal{F}_s]$. Here, Note $|E[t\wedge T-s|\mathcal{F}_s]-\Delta t|\leq\delta_X\Delta t$ where $\delta_X=P\{T<t\}$. In fact, since $t\wedge T=(s+\Delta t)\wedge T=s+\Delta t\wedge T$, $E[t\wedge T-s|\mathcal{F}_s]=E[\Delta t\wedge T|\mathcal{F}_s]\leq\Delta t$.

On the other hand,

$$
\begin{aligned}
E[\Delta t \wedge T|\mathcal{F}_s] &\geq \int_{\{T-s\geq t-s\}} \Delta t dP \\
&= \Delta t(1 - P\{T < t\}) \\
&= \Delta t(1-\delta_X).
\end{aligned}
$$

Thus, $0 \leq \Delta t - E[\Delta t \wedge T|\mathcal{F}_s] \leq \delta_X \Delta t$. We call this difference an error associated with stopping the process, X_t.

(2) Otherwise: Since $|\alpha+\beta-e_i| > 0$,

$$
\begin{aligned}
&E\left[\int_s^{t\wedge T} (X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\,\middle|\,\mathcal{F}_s\right] \\
&= E\left[\int_s^{t} (X_{i,u\wedge T}-X_{i,s})^{\alpha_i+\beta_i-1}\prod_{j\neq i}(X_{j,u\wedge T}-X_{j,s})^{\alpha_j+\beta_j}du\,\middle|\,\mathcal{F}_s\right] \\
&= \int_s^t \psi_{\alpha+\beta-e_i}(u)du.
\end{aligned}
$$

Because of $|\alpha+\beta-e_i| > 0$ the first equality holds. As a result, the main part is expressed as the sum of $f_i(X_s)\Delta t + f_i(X_s)(E[\Delta t \wedge T|\mathcal{F}_s] - \Delta t)$ and linear combination of $\int_s^t \psi_{\alpha+\beta-e_i}(u)du$ with $\mathcal{F}_s$–measurable coefficients $\partial^\beta f_i(X_s)/|\beta|!$. Next, we evaluate the remainder of the Taylor's expansion. Without loss of generality we can assume $|\partial^\beta f_i(X)| < K$ because $\tilde{X}_t$ is bounded and $f_i \in C^\infty(R^d)$. Define,

$$
R_{i,\beta} = E\left[\int_s^{t\wedge T}\sum_{|\beta|=n-m+2}\frac{\partial^\beta f_i(\xi_u)}{|\beta|!}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\,\middle|\,\mathcal{F}_s\right].
$$

Noticing $|\alpha+\beta-e_i| > 0$,

$$
\begin{aligned}
|R_{i,\beta}| &\leq E\left[\int_s^{t\wedge T}\left|\frac{\partial^\beta f_i(\xi_u)}{|\beta|!}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-1}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}\right|du\,\middle|\,\mathcal{F}_s\right] \\
&\leq \int_s^t \frac{K}{|\beta|!}\left(E\left[(X_{i,u\wedge T}-X_{i,s})^{2(\alpha_i+\beta_i-1)}\prod_{j\neq i}(X_{j,u\wedge T}-X_{j,s})^{2(\alpha_j+\beta_j)}\,\middle|\,\mathcal{F}_s\right]\right)^{1/2}du \\
&\leq \int_s^t \frac{K}{|\beta|!}(C_{n+1})^{1/2}(u-s)^{\frac{n+1}{2}}du \\
&= \frac{2K(C_{n+1})^{1/2}}{(n+3)(n-m+2)!}(\Delta t)^{\frac{n+3}{2}} \\
&< \frac{\sqrt{2}(2K^2 r\Delta t)^{\frac{n+3}{2}}\sqrt{(n+3)!}}{(n+3)(n-m+2)!}.
\end{aligned}
$$

The second inequality follows from Schwarz inequality and the third and fifth inequalities from the lemma 1. Now we define the last term as R^1_α.

Similarly, applying the Taylor's expansion of $g_{i,k}(X) = \sigma^2_{i,k}(X)$ up to $(n-m+3)$–th order to the second part, we get,

$$
\begin{aligned}
&E\left[\int_s^{t\wedge T}(X_{i,u}-X_{i,s})^{\alpha_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j}g_{i,k}(X_u)du\,\middle|\,\mathcal{F}_s\right]\\
&= \sum_{0\le|\beta|\le n-m+2}\frac{\partial^\beta g_{i,k}(X_s)}{|\beta|!}E\left[\int_s^{t\wedge T}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\,\middle|\,\mathcal{F}_s\right]\\
&\quad+\sum_{|\beta|=n-m+3}E\left[\frac{\partial^\beta g_{i,k}(\xi'_u)}{|\beta|!}\int_s^{t\wedge T}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\,\middle|\,\mathcal{F}_s\right].
\end{aligned}
$$

The conditional expectation in the main part of the Taylor's expansion is expressed as follows. Neglecting zero terms,

(1) $\alpha = 2e_i$ $(1 \le i \le d)$ and $|\beta| = 0$: Then, the conditional expectation is $E[\Delta t\wedge T - \Delta t|\mathcal{F}_s]$.
(2) Otherwise: Then, the conditional expectation is $\int_s^t \psi_{\alpha+\beta-2e_i}(u)du$.

In the same as the first part, the main part is expressed as the sum of $g_{i,k}(X_s)\Delta t + g_{i,k}(X_s)(E[\Delta t\wedge T|\mathcal{F}_s]-\Delta t)$ and linear combination of $\int_s^t \psi_{\alpha+\beta-2e_i}(u)du$ with $\mathcal{F}_s$–measurable coefficients $\partial^\beta g_{i,k}(X_s)/|\beta|!$. Evaluating the remainder of the Taylor's expansion, $R_{i,k,\beta}$, we get,

$$
\begin{aligned}
|R_{i,k,\beta}| &= \left|E\left[\int_s^t \frac{\partial^\beta g_{i,k}(\xi'_u)}{|\beta|!}(X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-2}\prod_{j\neq i}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j}du\,\middle|\,\mathcal{F}_s\right]\right|\\
&< \frac{\sqrt{2}(2K^2r\Delta t)^{\frac{n+3}{2}}\sqrt{(n+3)!}}{(n+3)(n-m+3)!},
\end{aligned}
$$

because $g_{i,k}\in C^\infty(R^d)$. We define the last term as R^2_α. Similarly, by setting $g_{i,j,k}(X) = \sigma_{i,k}(X)\sigma_{j,k}(X)$ we can show that

$$
E\left[\int_s^t (X_{i,u}-X_{i,s})^{\alpha_i-1}(X_{j,u}-X_{j,s})^{\alpha_j-1}\prod_{l\neq i,l\neq j}(X_{l,u}-X_{l,s})^{\alpha_l}g_{i,j,k}(X_u)du\right]
$$

can be expressed as the sum of the three parts. The first is $g_{i,j,k}(X_s)\Delta t + g_{i,j,k}(X_s)(E[\Delta t\wedge T|\mathcal{F}_s]-\Delta t)$, the second is linear combination of $\int_s^t \psi_{\alpha+\beta-e_i-e_j}(u)du$ with $\mathcal{F}_s$–measurable coefficients $\partial^\beta g_{i,j,k}(X_s)/|\beta|!$, and the last is the remainder of the Taylor's expansion, $R_{i,j,k,\beta}$, which satisfies,

$$\begin{aligned}
|R_{i,j,k,\beta}| &= \left| E\left[\int_s^t \frac{\partial^\beta g_{i,j,k}(\xi''_u)}{|\beta|!} (X_{i,u}-X_{i,s})^{\alpha_i+\beta_i-1}(X_{j,u}-X_{j,s})^{\alpha_j+\beta_j-1} \right.\right. \\
&\qquad \left.\left. \prod_{l\neq i,j}(X_{l,u}-X_{l,s})^{\alpha_l+\beta_l} du \middle| \mathcal{F}_s \right] \right| \\
&< R_\alpha^2.
\end{aligned}$$

Summing up three kinds of the remainders, we obtain the following total remainder associated with the expansion, $R_\alpha(t)$, for the conditional moments $\psi_\alpha(t)$:

$$\begin{aligned}
|R_\alpha(t)| &= \left| \sum_i \alpha_i \sum_{\beta=n-m+2} R_{i,\beta} \right. \\
&\quad \left. + \frac{1}{2}\left(\sum_{i,k}{}^2 \alpha_i(\alpha_i-1) \sum_{\beta=n-m+3} R_{i,k,\beta} + \sum_{i\neq j,k}{}^3 \alpha_i\alpha_j \sum_{\beta=n-m+3} R_{i,j,k,\beta} \right) \right| \\
&< \sum_i \alpha_i \sum_{\beta=n-m+2} R_\alpha^1 + \frac{1}{2}\left(\sum_{i,k}{}^2 \alpha_i(\alpha_i-1) \sum_{\beta=n-m+3} R_\alpha^2 + \sum_{i\neq j,k}{}^3 \alpha_i\alpha_j \sum_{\beta=n-m+3} R_\alpha^2 \right) \\
&= \left(d^{n-m+2} R_\alpha^1 + \frac{r}{2}\left(\left(\sum_i \alpha_i^2 - m \right) + \left(m^2 - \sum_i \alpha_i^2 \right) \right) d^{n-m+3} R_\alpha^2 \right) \\
&= \frac{\sqrt{2} m d^{n-m+2} \sqrt{(n+3)!}}{(n+3)(n-m+2)!} \left(1 + \frac{rd(m-1)}{2(n-m+3)} \right) (2K^2 r\Delta t)^{\frac{n+3}{2}}.
\end{aligned}$$

Similarly, for the total error associated with stopping the process, $R_{X,\alpha}(t)$,

$$\begin{aligned}
&|R_{X,\alpha}(t)| \\
&\quad < \sum_i \alpha_i \sum_{\beta=n-m+2} K\delta_X \Delta t \\
&\qquad + \frac{1}{2}\left(\sum_{i,k}{}^2 \alpha_i(\alpha_i-1) \sum_{\beta=n-m+3} K\delta_X\Delta t + \sum_{i\neq j,k}{}^3 \alpha_i\alpha_j \sum_{\beta=n-m+3} K\delta_X\Delta t \right) \\
&\quad = \left(d^{n-m+2} K\delta_X\Delta t + \frac{r}{2}\left(\left(\sum_i \alpha_i^2 - m \right) + \left(m^2 - \sum_i \alpha_i^2 \right) \right) d^{n-m+3} K\delta_X \Delta t \right) \\
&\quad = m d^{n-m+2} \left(1 + \frac{rd(m-1)}{2} \right) K\delta_X \Delta t.
\end{aligned}$$

This completes the proof.

Proof of Theorem 2. From theorem 1,

$$\Psi(t) = A\int_s^t \Psi(u)du + b(t) + R(t) + R_X(t).$$

Note that,

$$\frac{d}{dt}\exp(-At)\int_s^t \Psi(u)du = \exp(-At)\left(-A\int_s^t \Psi(u)du + \Psi(t)\right).$$

Then,

$$\Psi(t) = A\int_s^t \exp(A(t-u))(b(u) + R(u) + R_X(u))du + b(t) + R(t) + R_X(t).$$

We evaluate the remainder,

$$R_\Psi(t) = A\int_s^t \exp(A(t-u))(R(u) + R_X(u))du + R(t) + R_X(t).$$

From theorem 1 there exist constants C_R and C_X such that $|R(t)| < C_R(\Delta t)^{(n+3)/2}J$ and $|R_X(t)| < C_X(\Delta t)J$ where J is a column vector whose elements are all one. In fact we can choose them as follows,

$$\begin{aligned} C_R &= \frac{\sqrt{2}nd^{n+1}\sqrt{(n+3)!}}{2(n+3)}\left(1 + \frac{rd(n-1)}{6}\right)(2K^2r)^{\frac{n+3}{2}} \\ C_X &= nd^{n+1}\left(1 + \frac{rd(n-1)}{2}\right)K\delta_X. \end{aligned}$$

Then,

$$\begin{aligned} |R_\Psi| &\le \int_s^t |A|\exp(|A|(t-u))(|R(u)| + |R_X(u)|)du + |R(t)| + |R_X(t)| \\ &\le \int_s^t |A|\exp(|A|(t-u))du(C_R(\Delta t)^{\frac{n+3}{2}} + C_X(\Delta t))J + (C_R(\Delta t)^{\frac{n+3}{2}} + C_X(\Delta t))J \\ &= \exp(|A|\Delta t)(C_R(\Delta t)^{\frac{n+3}{2}} + C_X(\Delta t))J. \end{aligned}$$

where the inequality holds elementwise and every element of $|A|$ is its absolute value. Recall that for each α $(1 \le |\alpha| \le n)$, β $(0 \le |\beta| \le n - |\alpha| + 1)$ and e_i $(1 \le i \le d)$, the nonzero element of A is expressed as either of the following;

for $A_{\alpha,\alpha+\beta-e_i}$ ($|\alpha| \geq 2$ or $|\beta| \geq 1$),

$$\alpha_i \frac{\partial^\beta f_i(X_s)}{|\beta|!} + \sum_{\substack{j \neq i \\ \beta' = \beta - e_i + e_j}} \alpha_j \frac{\partial^{\beta'} f_j(X_s)}{|\beta'|!}$$

$$+ \sum_{\substack{j,k \\ \beta' = \beta - e_i + 2e_j}}^{2} \frac{\alpha_j(\alpha_j - 1)}{2} \frac{\partial^{\beta'} g_{j,k}(X_s)}{|\beta'|!}$$

$$+ \sum_{\substack{j \neq l,k \\ \beta' = \beta - e_i + e_j + e_l}}^{3} \frac{\alpha_j \alpha_l}{2} \frac{\partial^{\beta'} g_{j,l,k}(X_s)}{|\beta'|!},$$

or for $A_{\alpha,\alpha-2e_i}$ ($|\alpha| \geq 2$),

$$\sum_{i,k}^{2} \frac{\alpha_i(\alpha_i - 1)}{2} g_{i,k}(X_s) + \sum_{j \neq l,k}^{3} \frac{\alpha_j \alpha_l}{2} g_{j,l,k}(X_s).$$

For the first case, since all the partial derivatives are bounded by K and at least $|\beta| \geq 0$,

$$\begin{aligned} |A_{\alpha,\alpha+\beta-e_i}| &\leq \frac{|\alpha|K}{|\beta|!} + \frac{Kr}{2|\beta'|!}\left(\sum_i \alpha_i^2 - |\alpha|\right) + \frac{Kr}{2|\beta'|!}\left(|\alpha| - \sum_i \alpha_i^2\right) \\ &= \frac{|\alpha|K}{|\beta|!}\left(1 + \frac{r(|\alpha| - 1)}{2(|\beta| + 1)}\right) \\ &\leq nK\left(1 + \frac{r(n-1)}{2}\right). \end{aligned}$$

Clearly, the inequality holds for $A_{\alpha,\alpha-2e_i}$. Designating the last term as C_A, we get,

$$|A| \leq C_A H,$$

where H is a matrix whose elements are all one. Noticing J is an eigen vector for the eigen value, $\phi(n)$, of H,

$$\exp(C_A H \Delta t) J = \exp(C_A \phi(n) \Delta t) J.$$

Finally,

$$|R_\Psi| < \exp(C_A \phi(n) \Delta t)(C_R (\Delta t)^{\frac{n+3}{2}} + C_X(\Delta t)) J.$$

This completes the proof.

Proof of Corollary 1. The proof of theorem 2 shows that the remainder associated with the Taylor's expansion, $R(t)$, is zero-vector if $\partial^\beta f_i = 0$ ($|\beta| = n - m + 2$) and $\partial^\beta g_{i,k} = \partial^\beta g_{i,j,k} = 0$ ($|\beta| = n - m + 3$) for all m ($1 \leq m \leq n$), where $1 \leq i \neq j \leq d$ and $1 \leq k \leq r$. This holds when $\partial^\beta f_i = 0$ ($|\beta| = 2$) and $\partial^\beta g_{i,k} = \partial^\beta g_{i,j,k} = 0$ ($|\beta| = 3$). On the other hand, if $f_i(X)$ is linear in X and both $g_{i,k}(X)$ and $g_{i,j,k}(X)$ are at most quadratic, then, these equalities hold.

Moreover, focusing on the proof of theorem 2, boundedness of the process, X_t, is used only to evaluate $R(t)$. In this case $R(t) = 0$, so the boundedness is not required. This completes the proof.

References

[1] Bernstein, D. S. and So, W., Some Explicit Formulas for the Matrix Exponential, *IEEE Trans. Automat. Control*, **38** (1993), 1228–1232.

[2] Bibby, B. M. and Sorensen, M., Martingale Estimation Functions for Discretely Observed Diffusion Processes, *Bernoulli*, **1** (1995), 17–39.

[3] Biscay, R., Jimenez, J. C., Riera, J. and Valdes, P. A., Local Linearization Method for the Numerical Solution of Stochastic Differential Equations, *Annals of Institute Statistical Mathematics*, **48** (1996), 631–644.

[4] Chan, K. C., Karolyi, G. A., Longstaff, F. A. and Sanders, A. B., An Empirical Comparison of Alternative Models of the Short-Term Interest Rates, *J. Finance*, **47** (1992), 1209–1227.

[5] Cox, J. C., Ingersoll, J. E. and Ross, S. A., A Theory of the Term Structure of Interest Rates, *Econometrica*, **53** (1985), 385–408.

[6] Hansen, L. P., Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, **50** (1982), 1029–1054.

[7] Karatzas, I. and Shreve, S. E., *Brownian Motion and Stochastic Calculus. 2nd ed.*, Springer, New York, Berlin (1991).

[8] Kloeden, P. E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer, New York, Berlin (1992).

[9] Pedersen, A. R., A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations, *Scand. J. Statist.*, **22** (1995), 55–71.

[10] Prakasa Rao, B. L. S., *Statistical Inference for Diffusion Type Processes*, Arnold, London (1999).

[11] Shoji, I., Approximation of Continuous Time Stochastic Processes by a Local Linearization Method, *Math. Comp.*, **67** (1998), 287–298.

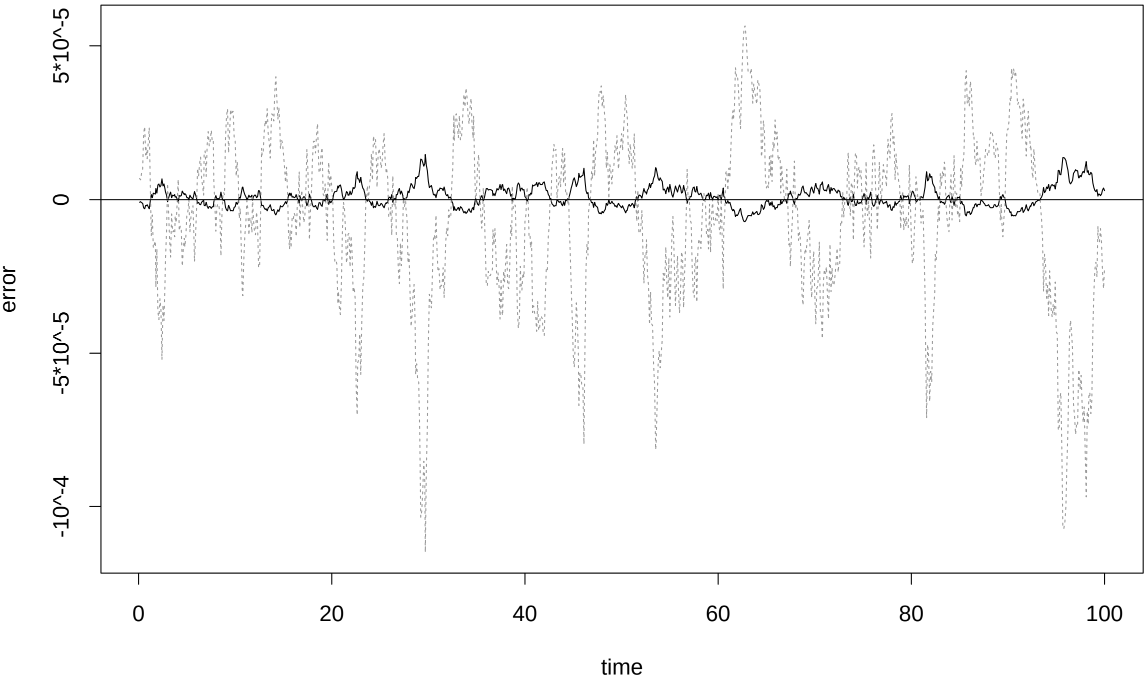
[12] White, H., Maximum Likelihood Estimation of Misspecified Models, *Econometrica*, **50** (1982), 1–25.

[13] White, H., *Estimation, Inference and Specification Analysis*, Cambridge University Press (1994).

[14] Wong, E., The Construction of a Class of Stationary Markov Processes, *Proc. Symposia in Appl. Math.*, **16** (1963), 264–276.

[15] Yoshida, N., Estimation for Diffusion Processes from Discrete Observation, *J. Multivariate Anal.*, **41** (1992), 220–242.

Figure 1
Comparison of Approximation Error: mean (upper) and variance (lower)



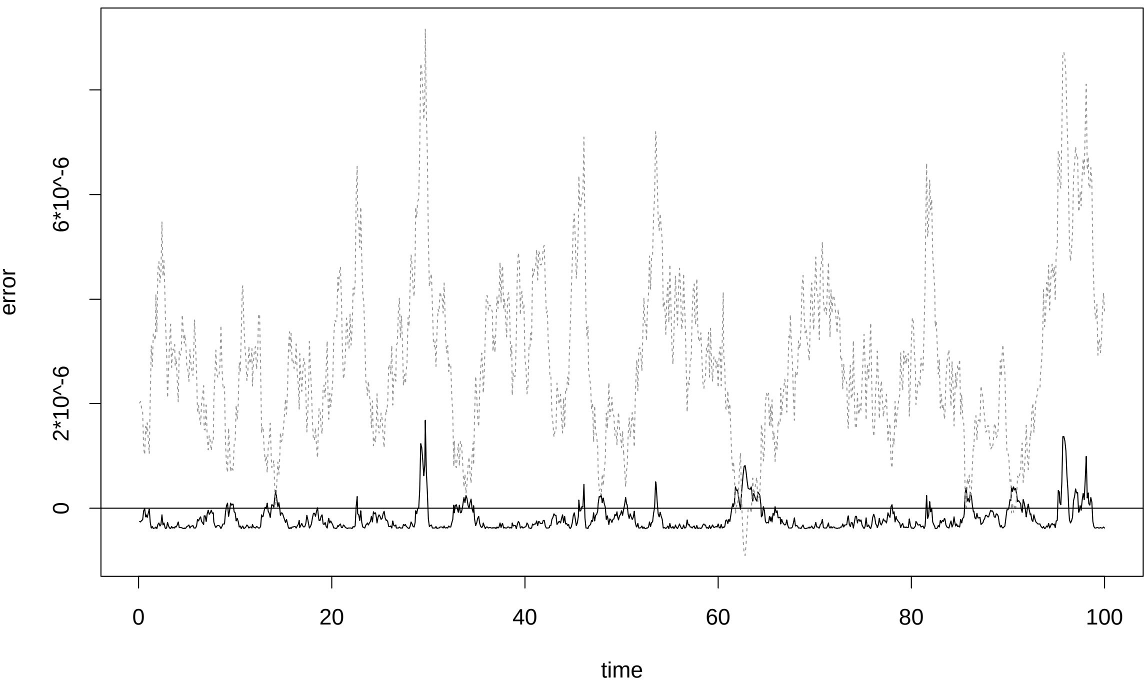
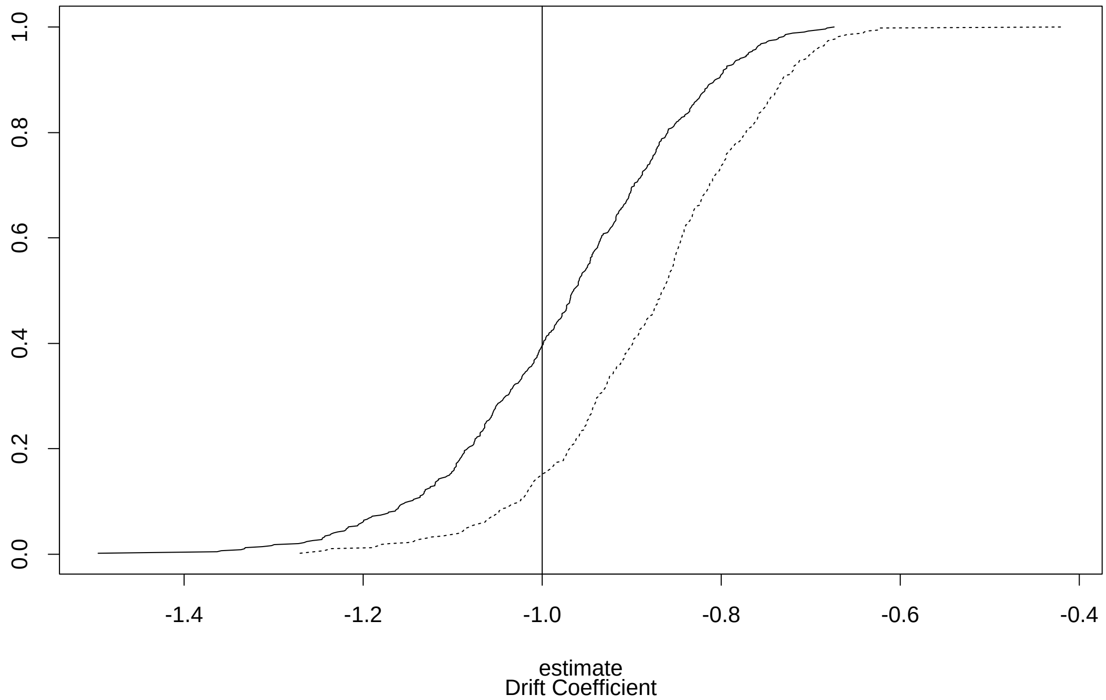
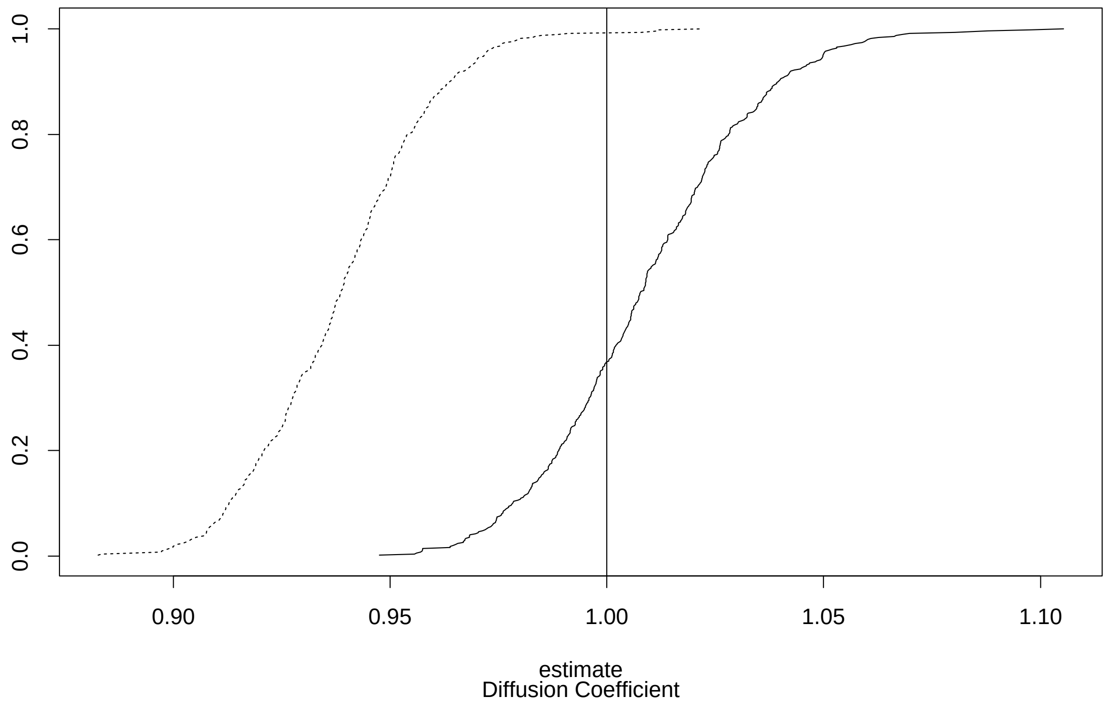


Figure 2



0.0
0.2
0.4
0.6
0.8
1.0
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
estimate
Drift Coefficient



0.0
0.2
0.4
0.6
0.8
1.0
0.90
0.95
1.00
1.05
1.10
estimate
Diffusion Coefficient