A test for multidimensional reducible diffusion models

Wang Jun[1]**, Chen Ping**[2]
Nanjing University of Science and Technology, Nanjing, China
[2]Corresponding author
E-mail: [1]*wangjun19901998@sina.com*, [2]*pancigm89@163.com*

Received 22 March 2020; accepted 9 April 2020
DOI https://doi.org/10.21595/vp.2020.21388

Check for updates

Copyright © 2020 Wang Jun, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. We develop a test for multidimensional diffusion models with known drift term based on transformation of diffusion matrix. The test in this article is different from the traditional method. We use a one-to-one transformation to transform the original model into a unit diffusion. Our approach is not only effective for stationary processes. On the other hand, our test performs well both on size and power.

Keywords: multidimensional diffusion models, reducible diffusion, one-to-one transformation, hypothesis and test.

1. Introduction

A random variable X_t satisfies the following stochastic differential equation and parametric specification model:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{1}$$
$$dX_t = \mu(X_t,\theta)dt + \sigma(X_t,\theta)dW_t, \tag{2}$$

where the random variable $X_t = \left(X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(1)}\right)$ is a d-dimension state vector on $S_X \subset R^d$, drift coefficient $\mu(X_t)$ and $\mu(X_t,\theta)$ are both d-dimension vector, diffusion coefficient $\sigma(X_t)$ and $\sigma(X_t,\theta)$ are both $d \times m$ matrix $m \leq d$, and W_t is a m-dimension standard Brownian motion.The focus of this paper is on testing the validity of the parametric specification Mode (2) based on a set of discretely observed data $\{X_{t\Delta}\}_{t=0}^n$.

For testing one-dimensional diffusion, in a pioneering work, Ait-Sahalia proposed an approach for testing the parametric specification model based on marginal density [1]. The advantage of the test is that the parametric marginal density of most of the diffusion processes is easy to know. But the method has several limitations. Hong and Li, Chen and Gao have an important development after Ait-Sahalia's work [2, 3]. Separately, the above two articles presents two different methods even though both of the methods are based on TPDF (transition probability distribution density). Chen shows that the method of Hong and Li will excessively accept the null hypothesis [4].

The work in multi-dimensional case is not going well. Hong and Li's method can be used in multidimensional case. But the result is not very satisfactory. Song develops a martingale approach [5]. "Martingale Problems" is used to transform null hypothesis several times. Then, the multi-dimensional problem is broken into multiple one-dimensional problems. Their method considers all relationship of any two variables, so the difficulty of calculation is geometric growth. This is the so-called dimensional disaster.

Based on Hermite series expansion, Ait-Sahalia makes a breakthrough in giving the closed-form of the approximate TPDF for a univariate time-homogeneous diffusion [6]. Ait-Sahalia extends his previous work to the multidimensional diffusions by using Kolmogorov equations [7]. On the basis of Ait-Sahalia, Choi proposes the approximate TPDF of multidimensional time-inhomogeneous diffusion models [8, 9]. In recent years, there are some other works on diffusion model testing [4, 10-12].

In this paper, we propose a new test method moved by Ait-Sahalia [7]. For most of diffusion processes, we can transfer diffusion X into a diffusion Y whose diffusion matrix σ_Y is the identity matrix by a one-to-one transformation. Then, we test if the diffusion matrix σ_Y is an identity matrix.

Our method has 3 advantages: (a) The condition of the strictly stationary process is not needed. Whether calculating the marginal density or the transitional density, the process needs to be strictly stationary process. While our method is adapted to non-stationary process. Only if the diffusion is a strictly stationary process, we can give the nonparametric estimator of TDPF. But our method does not need to calculate the nonparametric estimator of TDPF. (b) Our test improves the result of size and power. (c) It also works for more complicated cases. The most important improvement is dimensional disaster can be avoided.

The structure of the paper is as follows. In Section 2, we state the setup and assumptions. Section 3 reports the procedure and main results of the test. Simulation studies and empirical application are reported in the next two sections. At last, we conclude the discussion.

2. Setup and assumptions

Assumption 1. S_X is a product of m intervals with limits $\underline{x_i}$ and $\overline{x}_i$, where possibly $\underline{x_i} = -\infty$ and $\overline{x}_i = +\infty$, in which case, the intervals are open at infinite limits.

Assumption 2. $\forall x \in S_X$, the matrix $a(x) \equiv \sigma(x)\sigma'(x)$ is positive definite.

Assumption 3. For $i, j = 1,2, \dots, m$, $\mu_i(x)$ and $\sigma_{ij}(x)$ are infinitely differentiable with respect to $x \in S_X$ and $t \in [0, +\infty)$.

Assumption 4. There is a constant $K > 0$ such that for all $x \in [0, +\infty) \times S_X$:

$$\|\mu(x)\|^2 \leq K(1 + \|x\|^2), \quad \|\sigma(x)\|^2 \leq K(1 + \|x\|^2). \tag{3}$$

Assumption 3 and Assumption 4 ensure the uniqueness and existence of the solution of the Eq. (1) respectively. Indeed, Assumption 3 implies in particular that the coefficients of the stochastic differential equation are locally Lipschitz under their assumed (once) differentiability, which can be seen by applying the mean value theorem. And Assumption 4 can be relaxed in specific examples; it is not possible to do so in full generality.

3. Approach and test statistics

3.1. Reducible diffusions

Definition 1. If and only if there exists a one-to-one transformation of the diffusion X into a diffusion Y whose diffusion matrix σ_Y is the identity matrix, we can say that diffusion X is reducible. There exists an invertible function $\gamma(x)$ which is infinitely differentiable in X on S_X, such that $Y_t = \gamma(X_t)$ satisfies the stochastic differential equation on the domain S_X:

$$dY_t = \mu(Y_t)dt + dW_t. \tag{4}$$

If diffusion is reducible, the change of variable γ satisfies $\nabla\gamma(x) = \sigma^{-1}(x)$, by Ito's lemma. One-dimensional diffusion is reducible, by the simple transformation:

$$Y_t \equiv \nabla\gamma(X_t) = \int^{X_t} \frac{du}{\sigma(u)}, \tag{5}$$

where the lower bound of integration is an arbitrary point in the interior of S_X. The differentiability of γ ensures that μ_Y satisfies assumption 3. This change of variable is a Lamperti transform which plays a critical role in the derivation of closed-form Hermite approximations to the transition density of univariate diffusions. In next section, we give the method to solve the case that $1/\sigma(u)$ cannot be integrated in closed form. For multivariate case, not every diffusion is reducible. It depends on the specification of its σ matrix, in the following way.

Proposition 1. (Necessary and sufficient condition for reducibility). The diffusion X is said to be reducible if and only if:

$$\sum_{l=1}^{m}\frac{\partial\sigma_{ik}(x)}{\partial x_l}\sigma_{lj}(x)=\sum_{l=1}^{m}\frac{\partial\sigma_{ij}(x)}{\partial x_l}\sigma_{lk}(x), \tag{6}$$

for each x in S_X and triplet $x(i,j,k)=1,2,\dots,m$ such that $k>j$. If σ is non-singular, then the condition can be expressed as:

$$\frac{\partial[\sigma_{ij}^{-1}(x)]}{\partial x_k}=\frac{\partial[\sigma_{ik}^{-1}(x)]}{\partial x_j}. \tag{7}$$

Reducibility conditions are only for the matrix $\sigma(x)$. Under Proposition 1, when σ is nonsingular, $m^2(m-1)/2$ equalities must hold in order for an m-dimensional diffusion to be reducible. For example, when $m=2$, only two equalities need to be checked.

3.2. Irreducible diffusions

If the diffusion is irreducible, however, one no longer has the option of transforming X to Y. Notice that sample $\{X_1,X_2,\dots,X_n\}$ is discrete. Though we can't give the closed-form of transformation for diffusion X directly, we can transfer the sample $\{X_1,X_2,\dots,X_n\}$ to sample $\{Y_1,Y_2,\dots,Y_n\}$ of a unit diffusion Y discretely. For example, Ait-Sahalia points out that the diffusion is irreducible if σ is like case σ_{case1}:

$$\sigma_{case1}=\begin{bmatrix}\sigma_{11}(x_2) & 0\\ 0 & \sigma_{22}(x_2)\end{bmatrix},\qquad \sigma_{case2}=\begin{bmatrix}a(x_1)b(x_2) & a(x_1)c(x_2)\\ 0 & d(x_2)\end{bmatrix}. \tag{8}$$

Because σ_{11} depends on x_2. In practice, we treat the process of variable x_1 and x_2 in X_t as one-dimensional diffusion separately. So x_2 can be seen as a constant even though the x_2 changes every moment. Fortunately, the sample of x_2 is observed so that we can transform the sample of x_1 to the sample of a unit diffusion y_1.

If σ_{11} just depends on x_1, the diffusion process is reducible. Like σ_{case2}, we treat $b(x_2)$ in σ_{11} as constant. The matrix is similar to a diagonal matrix. Then it is like case 1.

Proposition 2. If irreducible diffusion X is diagonalized, we can transform the sample of diffusion X to a sample which belongs to a unit diffusion Y.

3.3. Hypothesis and test statistics

With known drift function, the hypothesis is as follows:

$$H_0: P[\sigma(X_t,\theta)=\sigma_0(X_t)]=1,\quad \exists\ \theta\in\Theta,\quad H_1: P[\sigma(X_t,\theta)=\sigma_0(X_t)]<1,\quad \forall\ \theta\in\Theta. \tag{9}$$

By Proposition 1 and Proposition 2, there exists a one-to-one transformation of a reducible diffusion to unit diffusion. So we just need to test whether the diffusion function is unit diffusion. We can transfer the reducible diffusion X to unit diffusion Y or the sample of irreducible diffusion X to sample of unit diffusion Y. The hypothesis becomes like the following:

$$H_0: P[\sigma(Y_t,\theta)=I]=1,\quad \exists\ \theta\in\Theta,\quad H_1: P[\sigma(Y_t,\theta)=I]<1,\quad \forall\ \theta\in\Theta, \tag{10}$$

where I is a unit matrix.

For the purpose, we need:

(1) Estimate the parameter θ of diffusion function by the observation sample data set $\{X_{t\Delta}\}_{t=0}^{n}$,
(2) Transfer $\{X_{t\Delta}\}_{t=0}^{n}$ with a unique transformation to sample $\{Y_{t\Delta}\}_{t=0}^{n}$ of a unit diffusion Y,
(3) Estimate the diffusion function of Y with $\{Y_{t\Delta}\}_{t=0}^{n}$.

Remark 1. In this paper, X_i means the i-th component in X. X_{t_i} means the observation at t_i- moment of X.

$\frac{n}{T}(Y_{t_{i+1}} - Y_{t_i})(Y_{t_{i+1}} - Y_{t_i})'$ can be treated as sample of $a(y)$. So our test statistics is:

$$T = \frac{n}{T}\left(Y_{t_{i+1}} - Y_{t_i}\right)\left(Y_{t_{i+1}} - Y_{t_i}\right)'. \tag{11}$$

Theorem 1. Under Assumptions 1-4, $T_{ii} \to \chi^2(1)$ and $T_{ij} \to 0$, $i \neq j$, under H_0.

Theorem 2. Under Assumptions 1-4, $T_{ii} \to \infty$ or $T_{ij} \to \infty$, $i \neq j$, under H_1.

Theorem 1 and Theorem 2 are applicable to one-dimensional situation.

4. Simulation experiment

4.1. Size evaluation

For examining the size of our test for multivariate models, we use a three-factor Vasicek model to generate the sample data. We also do size evaluation with Hong and Li's test and Song's test for comparison. Following [16], we set:

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0.1 & 0.2 & 2 \end{bmatrix}\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} dt + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} dW_t. \tag{12}$$

We use the asymptotic critical values (1.28 and 1.65) at the 10 and 5 % levels as the empirical rejection rates. Reference [8], we compute $T(\theta)$ over a bandwidth set $\mathcal{H} = \{h_k\}_{k=1}^{J}$. So the results in Table 1 are selected to be the result of the optimal bandwidth at the time of simulation.

Table 1. Sizes of $Q(j)$, Song's test and our $T(\theta)$ test

	$Q(j)$				Song				$T(\theta)$			
	X_{1t}	X_{2t}	X_{3t}	Total	X_{1t}	X_{2t}	X_{3t}	Total	X_{1t}	X_{2t}	X_{3t}	Total
5 %												
$n = 250$	4.5	3.4	2.9	3.7	4.6	5.1	4.9	6.1	4.3	4.3	3.3	3.4
$n = 500$	4.1	2.9	5.0	4.2	5.6	4.8	4.2	5.7	3.6	3.4	2.7	3.0
$n = 1000$	3.6	3.8	4.8	4.7	3.9	3.8	3.4	5.2	2.9	2.8	2.3	2.4
10 %												
$n = 250$	6.9	5.8	5.9	8.4	8.0	7.9	8.5	9.3	6.5	6.8	7.2	7.4
$n = 500$	7.0	6.5	6.9	8.2	7.7	7.6	7.8	8.6	5.8	6.3	6.6	6.9
$n = 1000$	7.3	7.1	6.7	7.7	7.0	6.5	7.2	7.7	5.3	5.2	6.1	6.2

Table 1 shows the result of sizes of three tests for three individual and combined generalized residuals separately at the 5 % and 10 % level. Overall, our test has a good performance on sizes at both 5 % and 10 % levels for sample sizes as small as $n = 250$ (i.e., about 20 years of monthly data). Our test makes an improvement to the size result.

Then, we consider the test for irreducible case. We use a bivariate Heston Model which is not a stationary process to generate 250, 500 and 1000 random samples. Table 2 shows that our test statistic also works on non-stationary process and has a good result. This is the most important breakthrough of our method. It makes $T(\theta)$ test can deal with more complex models:

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} rX_{1t} \\ b(a - X_{2t}) \end{bmatrix} dt + \begin{bmatrix} \sqrt{(1-\rho^2)X_{2t}}X_{1t} & \rho\sqrt{X_{2t}}X_{1t} \\ 0 & \sigma X_{2t} \end{bmatrix} dW_t. \tag{13}$$

Table 2. Sizes of test for Heston model

	5 %			10 %		
	$n = 250$	$n = 500$	$n = 1000$	$n = 250$	$n = 500$	$n = 1000$
X_{1t}	5.1	4.9	3.6	7.2	6.7	4.5
X_{2t}	6.4	5.7	3.9	8.1	6.5	5.0
Total	4.7	5.1	3.3	7.1	6.1	4.4

4.2. Power evaluation

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0.1 & 0.2 & 2 \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} dt + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} dW_t, \tag{14}$$

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0.1 & 0.2 & 2 \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} dt + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} dW_t. \tag{15}$$

Table 3 shows comparison result of $Q(j)$ our $T(\theta)$ test. We can see that, for the two models, we can determine which factors cause the error of null hypothesis, and from the whole, it is good to reject the null hypothesis. It will not accept the wrong null hypothesis too much, even though some components may be assumed to be correct. For either the $T(\theta)$ test or the $Q(j)$ test, the more the sample data n is, the larger the n is, the more significant the power is. For small samples. The power of the test needs to be improved, especially when $n = 250$. In general, power of our test is not as good as the size. Due to the principle of hypothesis testing, we can only ask for the size of our test to be as good as possible with finite sample.

Table 3. Power of $Q(j)$ our $T(\theta)$ test (in parentheses)

	X_{1t}	X_{2t}	X_{3t}	Total
Model 1				
n =250	37.2 (24.2)	6.3 (4.9)	7.7 (3.2)	17.2 (9.4)
$n = 500$	63.5 (50.8)	6.7 (4.8)	7.5 (5.2	27.2 (16.3)
$n = 1000$	94.6 (97.4)	6.6 (4.7)	6.6 (4.7)	53.7 (48.3)
Model 2				
$n = 250$	38.4 (26.7)	26.3 (13.7)	6.7 (4.5)	32.2 (18.2)
$n = 500$	74.2 (62.9)	48.1 (37.2)	7.2 (5.3)	65.7 (49.5)
$n = 1000$	95.3 (98.3)	78.5 (72.4)	6.8 (5.9)	93.2 (91.1)

5. Conclusions

In this article, we have developed a goodness of fit test for multidimensional diffusion models with known drift function. We use TPDF directly to construct our test statistic. Our method is effective for most multidimensional diffusion models. And it has a good performance in empirical applications. Of course, there is something to be improved. Even though our test works for most situations, there are still some cases that have not been covered. Next, we will focus on the test of drift function.

References

[1] **Ait-Sahalia Y.** Testing continuous-time models of the spot interest rate. Review of Financial Studies, Vol. 9, 1996, p. 385-426.

[2] **Hong Y., Li H.** Nonparametric specification testing for continuous-time models with application to spot interest rates. Review of Financial Studies, Vol. 18, 2005, p. 37-84.
[3] **Chen S. X., Gao J., Tang C. Y.** A test for model specification of diffusion processes. The Annals of Statistics, Vol. 36, 2008, p. 167-198.
[4] **Chen B., Song Z. G.** Testing whether the underlying continuous-time process follows diffusion: an infinitesimal operator-based approach. Journal of Econometrics, Vol. 173, 2013, p. 83-107.
[5] **Song Z. G.** A martingale approach for testing diffusion models based on infinitesimal operator. Journal of Econometrics, Vol. 162, Issue 2, 2011, p. 189-212.
[6] **Ait-Sahalia Y.** Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. Econometrica, Vol. 70, Issue 1, 2002, p. 223-262.
[7] **Ait-Sahalia Y.** Closed-form likelihood expansions for multivariate diffusions. The Annals of Statistics, Vol. 81, Issue 3, 2008, p. 637-654.
[8] **Choi S.** Closed-form likelihood expansions for multivariate time-inhomogeneous diffusions. Journal of Econometrics, Vol. 174, Issue 2, 2013, p. 45-65.
[9] **Choi S.** Explicit form of approximate transition probability density functions of diffusion processes. Journal of Econometrics, Vol. 187, Issue 1, 2015, p. 57-73.
[10] **Corradi V., Swanson N. R.** Predictive density construction and accuracy testing with multiple possibly misspecified diffusion models. Econometrics, Vol. 161, 2011, p. 304-324.
[11] **Negri I., Zhou L.** On goodness-of-fit testing for ergodic diffusion process with shift parameter. Stat Inference Stoch Process, Vol. 17, 2014, p. 51-73.
[12] **Zu Y.** Nonparametric specification tests for stochastic volatility models based on volatility density. Journal of Econometrics, Vol. 187, 2015, p. 323-344.