Interactive Estimator for Stochastic Differential Equation

Ryoji Fukuda
rfukuda@cc.oita-u.ac.jp
Faculty of Engineering,
Oita University,
Japan

Tatsumune Abe
abe@oz.ces.kyutech.ac.jp
Faculty of Computer Science and Systems Engineering,
Kyushu Institute of Technology,
Japan

Abstract: *Stochastic differential equations (SDE) are being applied widely; however, theory behind the concept is difficult to understand. Therefore, we designed an educational system for simple SDEs. The SDEs used in this system are determined by two linear functions with constant coefficients. Then, four real numbers are used to define this equation. In our system, a graph of a sample path with respect to an SDE is given, and the purpose is to estimate the four real numbers. The system suggests these values and some provides a few graphs of sample paths for an SDE corresponding to the given parameters. Using our system, a user should be able to understand the role of these SDE parameters.*

1. Introduction

A stochastic differential equation (SDE) expresses the changes in a stochastic process using the standard Lebesgue and the stochastic integrals. The short time average gradient and the corresponding variance are determined according to time and process value. SDEs are used in various probability theories; for example, the application to mathematical finance is well known. We designed an experimental virtual education system that helps in understanding SDEs. Our system use a linear SDE with constant coefficients. Both the Lebesgue and the stochastic integral terms are determined by using two real parameters, and then, four parameters are used to define this equation.

The target SDE is very simple, and the role of each parameter is well defined. Then, with a reasonably sufficient understanding of SDE graph shapes, one should be able to guess the parameters from the graph of a sample path of the solution. Our system displays the graph of one sample path of an SDE solution. The user does not know its parameters, and the purpose of the system is to estimate the parameters using the graph.

The SDE has a unique (strong) solution as a stochastic process [1]; however, the sample paths are different from each other according to the values of the Brownian motion process. In other words, two sample paths of the same solution are identical only if the corresponding values of Brownian motion, which controls the error values, are identical. Thus; it is not easy to judge whether two sample paths are controlled by the same SDE. Our system displays another sample path for which the four parameters are known. The values of this sample path are given in term of random numbers, and the shape of the graph changes every time a user draws the graph even if the four parameters remain unchanged. Moreover, the parameters are listed in a text box below the graphic area, and a user can change the parameters and generate sample path that corresponds to the new parameter values. By Iterating these tasks; the user may be able to estimate the parameters of the target sample path. The starting values for parameter estimation are generated automatically. In the case where the target sample path is generated by the system, the vector corresponding to the starting values is a neighborhood point in R^4 . Otherwise, or in the case where we assume that the parameters are unknown, the parameters are suggested using the values of the target sample path. An SDE approximates an increment in a process by using a linear combination of the increment in time and the increment in Brownian motion. Thus, it is possible to obtain approximated sequence of i.i.d. (independent identically distributed) N(0,1) random variables. According to some properties of an i.i.d. N(0,1) sequence, we prepared an evaluation function for target the SDE and estimated the starting values using this prepared function.

The first purpose of this estimation model was to evaluate drying shrinkage of concrete, however the above estimations are not correct enough and some new ideas are required to improve this estimation. We believe that human brains have an ability to recognize differences of SDE parameters by watching corresponding graphs, and expect that we can find some rules to find correct SDE parameters using our system. Then, for example, we will be able to find some estimation method for long term shrinkage value using short time data in near future.

2. Outline of the System

The purpose of our system is to estimate the (parameters of) SDE using a corresponding sample path. A comparative sample path is also displayed, for which the four parameters are known and this path is renewable. First suggestion of the parameters is given by the system, and after comparing these sample paths we can change the parameters.

2.1 Stochastic Differential Equation

In this study, we consider an SDE of the following type

$$X_t - X_s = \int_s^t (aX_r + b)dr + \int_s^t (c\,X_r + d)\,dB_r \quad , \tag{1}$$

where a, b, c, d are real constants, dr denotes the Lebesgue integral and dB_r denotes the Ito stochastic integral with respect to the Brownian motion process $B(t)$. This equation is often represented as

$$dX_t = (a\,X_t + b)\,dt + (c\,X_t + d)\,dBt \quad . \tag{2}$$

It is well known that equation (1) has a unique strong solution [1] and which is given by

$$X(t) = U(t)\left(X(0) + \int_0^t \frac{a - c*d}{U(s)}\,ds + \int_0^t \frac{c}{U(s)}\,dB(s)\right),$$

$$U(t) = \exp\left(\int_0^t \left(b - \frac{1}{2}d^2\right)ds + \int_0^t d\,dB(s)\right) \quad . \tag{3}$$

Using this solution and random numbers, we are able to obtain the values of the sample path of SDE (1) for the given parameters a, b, c, d.

2.2 Layout

The upper region of the system window, known as graphic area, is used to display the sample paths of the target SDE and the comparative SDE. Under the graphic area, there are text boxes for in putting the SDE parameters (a, b, c, d), buttons for drawing paths, and a text box for system messages.

3. Sample path of the solution

A stochastic process is random, however, it does not imply that the process is disordered. If we consider random events in Wiener space, two processes are identical strictly if the corresponding Brownian motions are identical. The solution of an SDE is a continuous-time stochastic process and natural phenomena occur continuously; usually we observe them as discrete time data. The low of a continuous-time stochastic process approximated by that of discrete time process. However, properties of the observed discrete process may depend on finer dividing stochastic processes. Therefore, we generate random variates as partial observations for the solution process as follows.

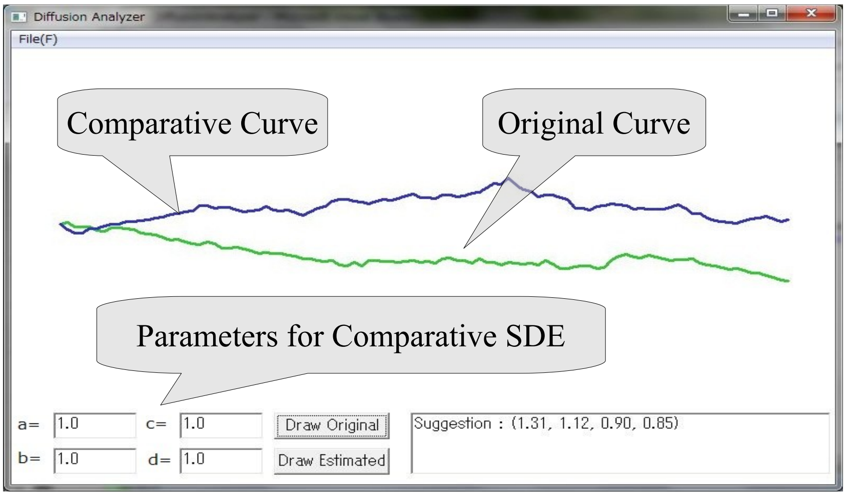

Figure 2.1 System Layout

1. Let $\{t_k\}_{k=0}^{n}$ be the dividing time of $[0,T]$ ($t_0=0, t_n=T$)
2. Let $\{s_{j+mk}\}_{j=0}^{m}$ be the equally spaced dividing time of $[t_{k-1}, t_k]$ ($k=1,2,\ldots,n$). Then $\{s_j\}_{j=0}^{N}$ ($N=nm$) is a fine dividing time of $[0,T]$ ($n=m=100$ is the standard value of the system).
3. Consider $\{g_j\}_{j=1}^{N}$ an i.i.d. $N(0,1)$ sequence of random numbers and set $w_j=\sqrt{\frac{T}{N}}\sum_{i=1}^{j} g_i$, which is a sample value of $B(\frac{jT}{N})$ ($j \le N$).
4. It is clear that $\int_0^{s_j}(b-\frac{1}{2}d^2)\,ds=(b-\frac{1}{2}d^2)s_j$ and $\int_0^{s_j} d\, dB(s)=w_j d$, then we obtain $U(s_j)$, $j \le N$ in (3).
5. We obtain approximated values of $\int_0^{t_k}\frac{a-c*d}{U(s)}ds$ ($j \le N$) using Simpson's rule on each $[t_{r-1}, t_r]$ ($r \le k$).
6. $\int_0^{t_k}\frac{c}{U(s)}dB(s)$ is approximated by using $\sum_{j=0}^{mk}\frac{c}{U(s_{j-1})}(w_j-w_{j-1})$.
7. We obtain the sample path $\{X(t_k)\}_{k=1}^{n}$ using the formula (3).

4. Measurement of SDE closeness

Assume that (2) is strictly true in case dt is replaced by t_k-t_{k-1} , dX_t by $X(t_k)-X(t_{k-1})$, and dB_t by $B(t_k)-B(t_{k-1})$. Then, we obtain the following equation.

$$\gamma_k=\frac{B(t_k)-B(t_{k-1})}{\sqrt{t_k-t_{k-1}}}=\frac{X(t_k)-X(t_{k-1})-(aX(t_{k-1})+b)(t_k-t_{k-1})}{(cX(t_{k-1})+d)\sqrt{t_k-t_{k-1}}}\sim N(0,1) . \quad (4)$$

$\{\gamma_k\}_{k=1}^{n}$ is the approximate i.i.d. $N(0,1)$ sequence because increments of a Brownian motion are independent. Note that the right hand side can be calculated using the process values, time values, and the SDE parameters.

The i.i.d. $N(0,1)$ sequence $\{\gamma_k\}_{k=1}^{n}$ satisfies the following conditions.

1. $\frac{1}{n}\sum_{k=1}^{n}\gamma_k \to 0$ (as $n\to\infty$). (5)
2. $\frac{1}{n/2}\sum_{k=1}^{n/2}(\gamma_{2k}-\bar{\gamma}_E)(\gamma_{2k-1}-\bar{\gamma}_O)\to 0$ (as $n\to\infty$). (6)
3. $\frac{1}{n/2}\sum_{k=1}^{n/2}(\gamma_k-\bar{\gamma}_F)^2-\frac{1}{n/2}\sum_{k=1}^{n/2}(\gamma_{k+n/2}-\bar{\gamma}_L)^2\to 0$ (as $n\to\infty$). (7)

where $\bar{\gamma}_E$, $\bar{\gamma}_O$, $\bar{\gamma}_F$, $\bar{\gamma}_L$ are the average values for even numbers, odd numbers, first half, and last half, respectively, of the sequence $\{\gamma_k\}_{k=1}^{n}$.

Using the Ito formula [1], for the process $Y(t)=X^2(t)$ we obtain:

$$dY(t)=A(t)dt+B(t)dWt$$
$$A(t)=2(aX_t+b)+(cX_t+d)^2$$
$$B(t)=2(cX_t+d)c_2$$

Then, by a similar calculation, we obtain:

$$\gamma'_k=\frac{dB_k}{\sqrt{dt_k}}=\frac{dY_k-A_k dt_k}{B_k\sqrt{dt_k}}\sim N(0,1) \tag{8}$$

where $dt_k=t_k-t_{k-1}, Y_k=Y(t_k)-Y(t_{k-1}), A_k=A(t_{k-1})$, and $B_k=B(t_{k-1})$. Then γ' (defined by (8)) satisfies similar properties with (5) $\sim$ (7) .

Using these properties we define the following functions to estimate closeness of SDE parameters.

$$\left.\begin{aligned} E_1&=\left(\frac{1}{n}\sum_{k=1}^{n}\gamma_k\right)^2\\ E_2&=\left|\frac{1}{n/2}\sum_{k=1}^{n/2}(\gamma_{2k}-\bar{\gamma}_E)(\gamma_{2k+1}-\bar{\gamma}_O)\right|\\ E_3&=\left|\frac{1}{n/2}\left(\sum_{k=1}^{n/2}(\gamma_k-\bar{\gamma}_F)^2-\sum_{k=n/2+1}^{n}(\gamma_k-\bar{\gamma}_L)^2\right)\right|\end{aligned}\right\} \tag{9}$$

We also define E'_1, E'_2 , and E'_3 by replacing γ_k with γ'_k .

These values may be small for large enough n only if the parameters are correct. The purpose of these values is to evaluate the closeness of the parameters; therefore, values of these parameters can be quite different from the correct values. In such a case, γ_k or γ'_k often takes extremely large (or small) values. This has a negative effect on our estimation. Therefore, we truncate these functions by using some constant $M>0$.

$$\tilde{\gamma}_k=\gamma_k\wedge M\vee(-M) \quad , \quad \tilde{\gamma}_k'=\gamma'_k\wedge M\vee(-M)$$

We define E_4 and E'_4 as numbers in the case where $|\gamma_k|$ or $|\gamma'_k|$ exceeds M.
(The default value is $M=4$.)
Figure 4.1 is a graph for these feature values. The SDE parameters are given by N(0,1) random numbers and fixed them. 30 comparative parameters are created using random number,the difference from the corresponding SDE parameter is N(0,0.1) random number. The horizontal measure is the difference between the SDE and the comparison parameters. Each feature values never increase in general and the variations are not small, therefore we are not able to judge the

closeness of parameters by one feature value.

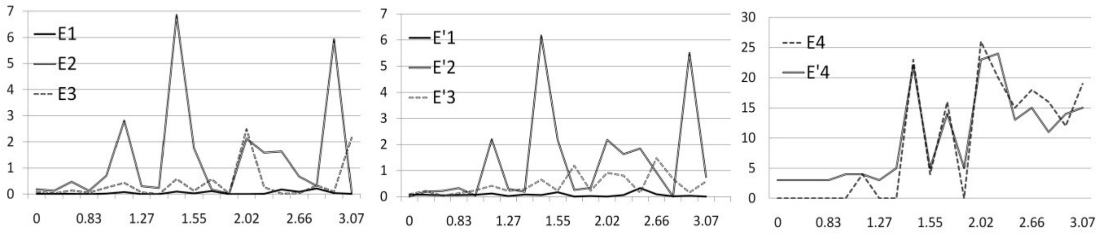

Figure 4.1 Feature values

5. Preliminary Estimation

To estimate the four parameters of the SDE, we use a simple generic algorithm. In this section, we explain our preliminary estimations. In the following estimations, we use some constants defined in several trials-and-errors.

5.1 Preliminary Estimation for the Legesgue integral part

For the preliminary estimation for a and b, we ignore the stochastic integral part. In the case where $c=d=0$ in (1) or (2), The equation becomes a standard differential equation and the solution is given as

$$f(x)=K e^{ax}-\frac{b}{a}, \text{ (} K \text{ is a constant).}$$

There are three parameters, namely, K, b, and a, in this function. The system finds the least square error solution for these parameters.

If the parameter a is a fixed constant, this estimation is a 2-dimensional linear regression, and for a fixed b and K, a can be approximated using the one-dimensional Newton method. Then, we approximate these values as follows:

1. The initial value of a is obtained using 3 pairs of time and process values.
2. Fix a, and find optimal values of b and K using the least square error method.
3. Next, fix b and K, and find a stationary point of the square error as a function of a using the Newton method.
4. Iterate steps 2. to 3. several times.

The default iteration number is 30. Without a stochastic integral term, the estimated values are correct enough (the errors are less than 10^{-3} when the target parameters are $1 \sim 3$). However, the actual shapes of the curves are quite different from ideal (non random) one, and we can not expect accurate estimation.

5.2 Preliminary Estimation for Stochastic Integral Part

We assume that the variance of the stochastic integral part is approximated by the least square error in the quadratic regression. Then, we estimate c and d by the following procedure. In the following explanation we use the same notations for X and $\{t_k\}$ ($k=1,2,..,n$) as those used in Sections 2 and 3.

1. Let $0<\rho<n$ be a positive integer (the standard value is 5). sd_k denotes the standard deviation of $\{X(t_j)-X(t_{j-1})-(L(t_j)-L(t_{j-1}))\}_{k-\rho\leq j\leq k+\rho}$ for each $k\leq n$, where L(t) is the corresponding regression line.

2. Ignore the indices if $j \le 0$ or $j > n$.
3. By the linear regression we obtain the approximation $sd_k \approx c\,X(t_k) + d$.

6. Suggestions of Parameters

We define an evaluation function of a set of parameters, which describes the closeness to the SDE parameters. Using this function we estimate the parameters. This estimation is not accurate enough and the graph shapes of sample paths are quite different from original one in general. However, we can not start interactive estimations without any information for parameters. Then the main purpose of this estimation is to find starting values for interactive estimations.

6.1 Generalized Choquet Integral with respect to a Two Additive Measure

The evaluation functions are defined as generalized Choquet integrals with respect to a two-additive measure [3].

Let A be an n-points set (int this case, $n=8$, i.e., the number of feature functions) and μ be a fuzzy measure or a general set function defined on the power set of A with $\mu(\emptyset)=0$. If μ is an additive measure, all measure values determined from $\mu(\{x\})$ ($x \in A$), that is, μ is defined by n values. However, a general set function is defined by 2^n values. Then, to reduce the number of parameters, we consider a two-additive measure as the first generalization of a non-additive measure analysis.

A two additive measure is given by

$$\mu(A) = \sum_{x \in A} \mu_x + \sum_{\{x,y\} \subset A} \nu_{x,y},$$

using real constants $\mu_x, \nu_{x,y}$ ($x, y \in A$). Then two-additive measure on an n-point set is defined by $n + n(n-1)/2$ values. The generalized Choquet integral of a function f ($A \to [0,1]$) is given by

$$\int_A f\, d\mu = \sum_{x \in A} f(x)\mu_x + \sum_{\{x,y\} \subset A} f(x) \otimes f(y)\, \nu_{x,y},$$

where μ is a two additive measure determined by $\mu_x, \nu_{x,y}$ ($x, y \in A$) and $\otimes$ is a Dombi t-norm [4] defined by

$$x \otimes y = \frac{1}{((1/x-1)^{\lambda} + (1/y-1)^{\lambda})^{1/\lambda}},$$

(in this case $\lambda = 2.5$, see [3]).

A (generalized) Choquet integral is an extension of a weighted sum. For an element $x \in A$, $f(x)$ has a large effect on the integral value when $|\mu_x|$ is large, and for a pair $x, y \in A$, in the case where $\nu_{x,y}$ is large, the integral value becomes large if both $f(x)$ and $f(y)$ take large values. Then, if we are able to select optimal values for these parameters in two additive measure, we only need to minimize this function to find the SDE parameters.

6.2 Estimation for a Two-Additive Measure

The purpose here is to create an optimal two-additive measure using the Monte-Carlo method. We obtain SDE parameters using random numbers and create comparative parameters using random numbers. We adopt the Euclidean distance in $\mathbb{R}^4$ as a criterion value for the regression. Then, we estimate the two-additive measure through the following steps.

1. Create 10 sets of parameters using N(0,1) random numbers. These are SDE parameters for observed stochastic processes.
2. For each set of parameters, we create 100 comparison parameters by adding N(0,0.1) random variables to the parameters given in the above step. Then, obtain the Euclidean distance in $\mathbb{R}^4$.
3. For each set of parameters given in the first step, create a sample path by the same method as that described in section 3.
4. For each sample path of the process and for each comparison parameter given in the second step, we obtain feature values $E_1 \sim E_4$ and $E'_1 \sim E'_4$.
5. Using the vectors of normalized evaluation values $\{\vec{e}_i = (e_i^{(j)})_{j=1,..,8}\}_{i=1}^{1000}$ and distance values $\{d_i\}_{i=1}^{1000}$, we approximate the two-additive measure μ to minimize $\sum_{i=1}^{100} (\int \vec{e}_i d\mu - d_i)^2 = \sum_{i=1}^{100} (\sum_j e_i^{(j)} \mu_j + \sum_{j,j'} e_i^{(j)} \otimes e_i^{(j')} \nu_{j,j'} - d_i)^2$, using $34(=8+{}_8C_2)$ dimensional linear regression.

6.3 Estimation for SDE Parameters

Using the two-additive measure given in the previous subsection, we estimate the SDE parameters by the following generic algorithm. In the estimation, we use 100 four- dimensional gene vectors (SDE parameters).

1. Initial values for the SDE parameters are given using the methods explained in Section 5. This vector is copied to all 100 gene vectors.
2. N(0,0.1) (independent) random numbers are added to every coefficient of all gene vectors.
3. The gene vectors are sorted in an ascending order of the evaluation function given in the previous subsection.
4. Recombination (we will explain this step below) is performed according to the sort result.
5. Estimated parameters are obtained after 50 iteration of 3. ~ 4.

The First three gene vectors are left intact. The remaining gene vectors are divided into three groups of the same size, and replaced by the following vectors:

0. Let $\{\vec{v}_i\}_{i=1}^{32}$ be the first 32 gene vectors, including the first 3 vectors.
1. Genes in the first group are replaced by improved vectors using the Newton method.
2. Genes in the second group are replaced by improved vectors using the gradient method. We use a random variable as the coefficient of the gradient.
3. Genes in the third group are replaced by $\{\vec{v}\}_{i=1}^{32}$, and then these values are modified using N(0,0.1) random numbers.

6.4 Samples of the Estimated Curve

Three original and estimated lines are listed in Figure 6.1. The user will find more adequate parameters after several trials-and-errors.

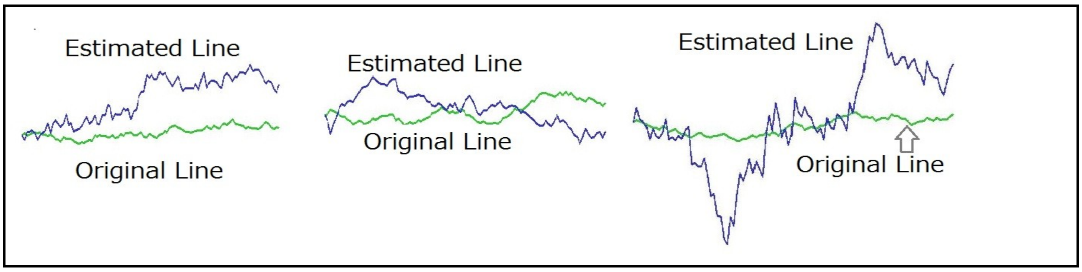

Figure 6.1 Original and Estimated Lines.

7. Conclusion

We designed an interactive estimator for SDE parameters. We defined eight feature functions for the correctness of the parameters and created an evaluation function as a Choquet integral with respect to a two-additive measure, where the two-additive measure was obtained by a Monte-Carlo simulation. Suggested values for SDE parameters for given data are given using this evaluation function.

Students can understand the roles played by the SDE parameters, and in the near feature, they can suggest solutions for several problem corresponding to SDEs by using our system.

References

[1] B. Oksendal, Stochastic Differential Equations: An Introduction with Applications, Springer, 1985

[2] T. Mikosch, Elementary Stochastic Calculus With Finance in View , Advanced Series on Statistical Science & Applied Probability, Vol 6 (1999)

[3] R. Fukuda, K. Nii (2009), Handwriting Tool for Note Takers in Mathematical Classes, Proceedings of the 14th Asian Technology Conference in Mathematics.

[4] J. Dombi, A General Class of Fuzzy operators, the DeMorgan Class of Fuzzy Operators and Fuzziness Measures Induced by Fuzzy Operators, Fuzzy Sets and Systems, 8, 149-163, 1982