УДК 681.513.54

ОЦЕНКА КОЭФФИЦИЕНТА ДИФФУЗИИ ВИНЕРОВСКОГО СЛУЧАЙНОГО ПРОЦЕССА С РАВНОМЕРНЫМ СНОСОМ

А.И. Саичев, В.А. Филимонов, М.В. Тараканова

Нижегородский госуниверситет им. Н.И. Лобачевского

martenok86@mail.ru

Поступила в редакцию 02.04.2010

Предлагаются два метода оценки коэффициента диффузии случайного винеровского процесса с равномерным сносом: метод наибольшего правдоподобия и метод наименьших квадратов. Показана несмещенность и найдены дисперсии полученных оценок, также представлено сравнение оценок, полученных данными двумя методами.

Ключевые слова: статистическая радиофизика, коэффициент диффузии, статистическая оценка параметров.

Введение

Одной из важнейших задач прикладной радиофизики является построение оптимальных оценочных формул различных измерений. Если объектом измерений является случайный процесс, как правило, представляется возможным получить лишь несколько реализаций процесса, а иногда лишь некоторые отсчеты внутри одной реализации. В этом случае оценочные формулы содержат максимум информации о параметрах измеряемого случайного процесса. Критериями качества полученных оценок являются несмещенность и эффективность. Несмещенными называют оценки параметров, у которых статистическое среднее полученной оценки равняется исходному параметру. Эффективными же являются оценки с минимальной дисперсией [1].

В данной статье рассматривается оценка параметров винеровского случайного процесса. Винеровским называется нестационарный случайный процесс с независимыми приращениями, распределенными по нормальному закону с нулевым средним и дисперсией $\langle dW^2\rangle \sim dt$. Этот процесс, являясь математической моделью броуновского движения или случайного блуждания с непрерывным временем, широко применяется в различных областях знаний: физике и геофизике, медицине и биофизике, химии и радиотехнике, а также в области передачи информации. В настоящее время возросший интерес к данному процессу обусловлен тем, что на его основе строится большинство моделей поведения различных финансовых показателей [2]. В общем случае исследуемый процесс может быть представлен в следующем виде:

$$X(t) = ct + \sqrt{D}W(t), \qquad (1)$$

где c – постоянная скорость сноса процесса, D – коэффициент диффузии, а $W(t)$ – винеровский процесс, заданный своими начальным и средним значениями, а также автокорреляционной функцией:

$$W(0)=0,\ \langle W(t)\rangle=0,\ \langle W(t_1)W(t_2)\rangle=\min(t_1,t_2).$$

Задачей, поставленной в данной работе, является оценка коэффициента диффузии D процесса (1), при условии, что известны отсчеты процесса x_i, взятые через равные промежутки времени $t_i = iT/n$, где T – длительность реализации, а n – количество отсчетов. В работе оценка проводится двумя различными методами: методом наибольшего правдоподобия и методом наименьших квадратов. Также проведены проверка полученных оценок на соответствие критериям качества и сравнение полученных результатов.

Метод наименьших квадратов

Метод наименьших квадратов – статистический метод определения неизвестных параметров случайных процессов путем минимизации среднеквадратичных отклонений между фактическими и расчетными данными [3].

Рассматривая винеровский процесс с равномерным сносом, представим среднеквадратичные отклонения для отыскания оценок скорости сноса $\hat{c}$ и коэффициента диффузии $\hat{D}$ в виде:

$$f_1(x_1,...,x_n)=\sum_{i=1}^{n}\left(x_i-\hat{c}t_i\right)^2 \to \min_{\hat{c}}, \qquad (2)$$

$$f_2(x_1,...,x_n)=\sum_{i=1}^{n}\left(\left(x_i-\hat{c}t_i\right)^2-\hat{D}T\right)^2 \to \min_{\hat{D}}. \qquad (3)$$

Найдем минимумы функций (2) и (3), для чего приравняем нулю частные производные по $\hat{c}$ и $\hat{D}$ соответственно:

$$\begin{cases} \dfrac{\partial f_1(x_1,...,x_n)}{\partial \hat{c}}=0 \\ \dfrac{\partial f_2(x_1,...,x_n)}{\partial \hat{D}}=0 \end{cases}.$$

Непосредственное решение приведенной системы уравнений приводит к следующим выражениям для оценок:

$$\hat{c}=\frac{1}{T}\frac{6}{(n+1)(2n+1)}\sum_{i=1}^{n} ix_i\ , \qquad (4)$$

$$\hat{D}=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\hat{c}\frac{T}{n}i\right)^2 . \qquad (5)$$

Проверим, отвечают ли полученные выражения критериям качества (оптимальности) статистических оценок, а именно несмещенности и эффективности. Проверим сначала оценку скорости сноса $\hat{c}$, поскольку эта оценка непосредственно входит в выражение (5) для оценки коэффициента диффузии.

Смещенность оценки определяется ее средним:

$$\left\langle \hat{c}\right\rangle=\frac{1}{T}\frac{6}{(n+1)(2n+1)}\sum_{i=1}^{n} i\left\langle x_i\right\rangle . \qquad (6)$$

Воспользовавшись средним значением i-го отсчета

$$\left\langle x_i\right\rangle=c\frac{T}{n}i\,,$$

получим

$$\left\langle \hat{c}\right\rangle=\frac{1}{T}\frac{6}{(n+1)(2n+1)}c\frac{T}{n}\sum_{i=1}^{n} i^2=c. \qquad (7)$$

Таким образом, полученная оценка скорости сноса является несмещенной. Вычислим дисперсию скорости сноса

$$Var\left[\hat{c}\right]=\left\langle \hat{c}^2\right\rangle-\left\langle \hat{c}\right\rangle^2 , \qquad (8)$$

где

$$\left\langle \hat{c}^2\right\rangle=\frac{1}{T^2}\left(\frac{6}{(n+1)(2n+1)}\right)^2\sum_{i,j=1}^{n} ij\left\langle x_i x_j\right\rangle . \qquad (9)$$

Воспользовавшись выражением для корреляции двух отсчетов процесса

$$\left\langle x_i x_j\right\rangle=D\frac{T}{n}\min\{i,j\}+\left(\frac{cT}{n}\right)^2 ij \qquad (10)$$

и тем, что

$$\sum_{i,j=1}^{n} j\min\{ij\}= =\frac{1}{30}n(n+1)(2n+1)(2n^2+2n+1)\,, \qquad (11)$$

из (9) получим

$$\left\langle \hat{c}^2\right\rangle=c^2+\frac{D}{T}\frac{6}{5}\frac{2n^2+2n+1}{(n+1)(2n+1)} . \qquad (12)$$

Дисперсия же оценки скорости сноса равна:

$$Var\left[\hat{c}\right]=\frac{6}{5}\frac{2n^2+2n+1}{2n^2+3n+1}\frac{D}{T} . \qquad (13)$$

Из выражения (13) видно, что дисперсия оценки растет с увеличением количества точек (рис. 1), асимптотически приближаясь к значению $6D/5T$ при $n\to\infty$.

Найдем среднее $\langle \hat{D}\rangle$ коэффициента диффузии

$$\left\langle \hat{D}\right\rangle=\frac{1}{nT}\sum_{i=1}^{n}\left\langle x_i^2\right\rangle- -\frac{(n+1)(2n+1)}{6T}\left\langle \hat{c}^2\right\rangle\left(\frac{T}{n}\right)^2 . \qquad (14)$$

После подстановки (10) и (12) в (14) получим среднее значение коэффициента диффузии, не равное оцениваемому параметру, а следовательно, смещенную оценку

$$\left\langle \hat{D}\right\rangle=\frac{(n-1)(n+1)}{10n^2}D . \qquad (15)$$

Несмещенная оценка $\langle \hat{D}_0\rangle$ получится после нормировки оценки на ее среднее:

$$\hat{D}_0=\frac{10n^2}{(n-1)(n+1)}\hat{D}= =\frac{10n}{(n-1)(n+1)}\sum_{i=1}^{n}\left(x_i-\hat{c}\frac{T}{n}i\right)^2 , \qquad (16)$$

очевидно, что в этом случае

$$\left\langle \hat{D}_0\right\rangle=D . \qquad (17)$$

Рассмотрим эффективность нормированной оценки коэффициента диффузии

$$Var\left[\hat{D}_0\right]=\left\langle \hat{D}_0^2\right\rangle-\left\langle \hat{D}_0\right\rangle^2= =\left(\frac{10n^2}{(n-1)(n+2)}\right)^2 Var\left[\hat{D}\right]. \qquad (18)$$

Дисперсия смещенной оценки коэффициента диффузии была найдена при помощи программного пакета Wolfram Mathematica

$$Var\left[\hat{D}\right] = \frac{(n-1)(n+2)(n^2+n+8)}{175n^4} D^2 . \quad (19)$$

Подставив (19) в (18), получим дисперсию нормированной оценки

$$Var\left[\hat{D}_0\right] = \frac{4}{7} \frac{n^2+n+8}{(n-1)(n+2)} D^2 . \quad (20)$$

В отличие от оценки скорости сноса, эффективность оценки коэффициента диффузии увеличивается с увеличением числа измеренных отсчетов реализации (рис. 2). Оценка коэффициента диффузии (20) при этом стремится к $\frac{4}{7} D^2$.

Метод наибольшего правдоподобия

Метод наибольшего правдоподобия – метод, позволяющий получить оценку неизвестных параметров из условия максимума функции правдоподобия. Как правило, в качестве функции правдоподобия выступает плотность вероятности или некоторая монотонно возрастающая функция от плотности вероятности. При работе с гауссовыми процессами обычно используют логарифм плотности вероятности [4].

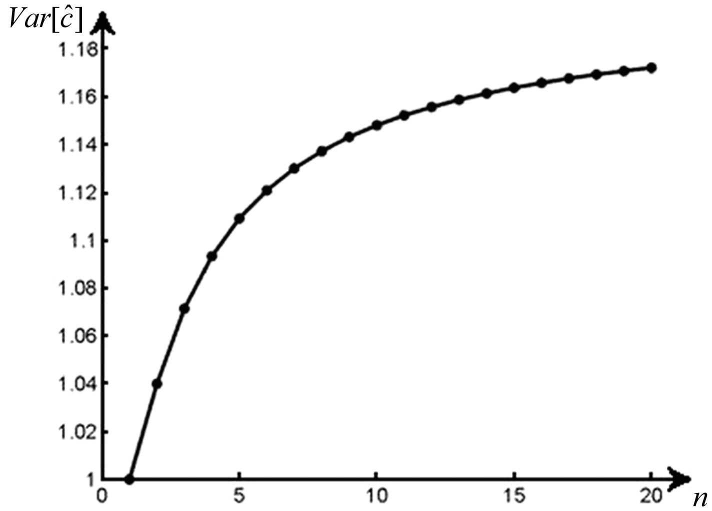

Рис. 1. Зависимость дисперсии оценки скорости сноса, полученной методом наименьших квадратов, от количества отсчетов n

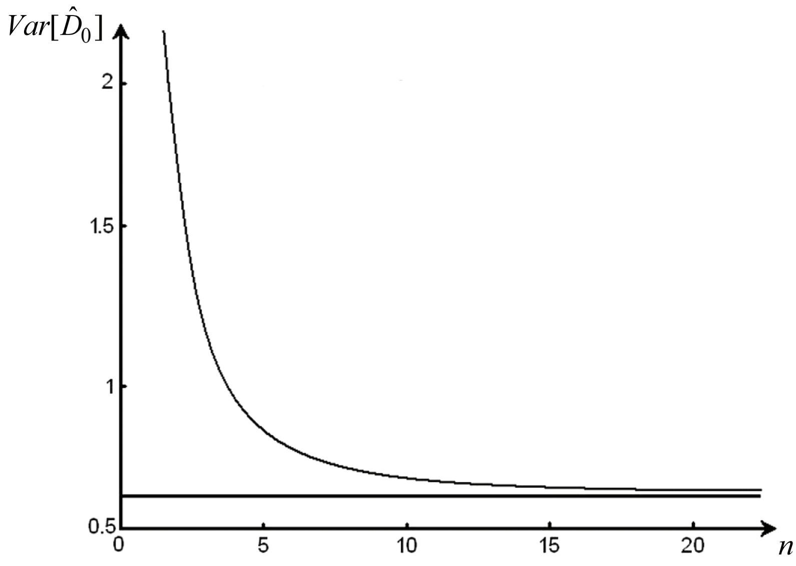

Рис. 2. Зависимость дисперсии оценки коэффициента диффузии, полученной методом наименьших квадратов, от количества отсчетов n. Горизонтальной линией показана асимптотика дисперсии оценки при $n \to \infty$

Для применения данного метода необходимо найти плотность вероятности винеровского случайного процесса с равномерным сносом.

Поскольку винеровский процесс является гауссовым, то его плотность вероятности имеет вид [5]:

$$f(x_1,t_1;...;x_n,t_n\,|\,c,D)= \\ =\frac{1}{\sqrt{(2\pi)^n\det(B_x)}}\exp\{-\frac{1}{2}\vec{x}^T B_x^{-1}\vec{x}\}, \qquad (21)$$

где $\vec{x}=x_i-ct_i$. Ковариационная матрица рассматриваемого процесса имеет вид:

$$B_x\left[t_i,t_j\right]=\left\langle(x_i-\langle x_i\rangle)\left(x_j-\langle x_j\rangle\right)\right\rangle, \qquad (22)$$

или

$$B_x\left[t_i,t_j\right]=\frac{DT}{n}\min(i,j). \qquad (23)$$

Обратная матрица

$$B_x^{-1}=\frac{n}{DT}\begin{pmatrix} 2 & -1 & 0 & ... & 0\\ -1 & 2 & -2 & ... & 0\\ 0 & -1 & 2 & ... & 0\\ ... & ... & ... & ... & ...\\ 0 & 0 & 0 & ... & 1\end{pmatrix}. \qquad (24)$$

Выполнив элементарные преобразования и сделав обратную замену $\widetilde{x}_i=x_i-ct_i$, из (21) получим выражение для плотности вероятности винеровского случайного процесса с равномерным сносом:

$$f(x_1,t_1,...,x_n,t_n\,|\,c,D)= \\ =\frac{1}{\sqrt{(2\pi)^n\left(\frac{DT}{n}\right)^n}}\exp\left\{-\frac{n}{DT}K(x_1,...,x_n,T,c)\right\}, \qquad (25)$$

где

$$K(x_1,...,x_n,T,c)= \\ =c^2T^2\frac{1}{2n}-cT\frac{x_n}{n}+\left(\sum_{i=1}^{n-1}x_i^2+\frac{x_n^2}{2}-\sum_{i=1}^{n-1}x_ix_{i+1}\right). \qquad (26)$$

Логарифмируя (25) придем к функции правдоподобия

$$\ln f(x_1,t_1,...,x_n,t_n\,|\,c,D)= \\ =\frac{n}{2}\ln\frac{n}{2\pi DT}+ \\ +\left\{-\frac{n}{DT}K(x_1,...,x_n,T,c)\right\}.$$

Выражения для оценок коэффициента диффузии и скорости сноса получаются из условия максимума функции правдоподобия. Для этого найдем частную производную от функции правдоподобия по переменным $c=\hat{c}$ и $D=\hat{D}$

$$\left.\frac{\partial\ln f(x_1,t_1;...;x_n,t_n\,|\,c,D)}{\partial c}\right|_{c=\hat{c},D=\hat{D}}= \\ =\left(2\hat{c}T^2\frac{1}{2n}-T\frac{1}{n}x_n\right)\left(-\frac{n}{\hat{D}T}\right), \\ \left.\frac{\partial\ln f(x_1,t_1;...;x_n,t_n\,|\,c,D)}{\partial D}\right|_{c=\hat{c},D=\hat{D}}= \\ =-\frac{n}{2}\frac{1}{\hat{D}}+\frac{n}{T}\frac{1}{\hat{D}^2}K(x_1,...,x_n,T,\hat{c}). \qquad (27)$$

Приравнивая выражения (27) к нулю, получим значения оценок параметров c и D

$$\hat{c}=\frac{x_n}{T}, \qquad (28)$$

$$\hat{D}=\frac{2K(x_1,...,x_n,T,\hat{c})}{T}. \qquad (29)$$

Рассмотрим, насколько полученные оценки являются эффективными, и проверим их несмещенность.

Результирующая оценка скорости сноса совпадает с оценкой, полученной методом наименьших квадратов. Поэтому значения среднего и дисперсии оценки скорости сноса совпадают с (7) и (13).

Для того чтобы определить, является ли оценка коэффициента диффузии смещенной и эффективной, представим ее в более удобном виде, подставив (26) в (29):

$$\hat{D}=\frac{2}{T}\left(\frac{x_n^2}{2}\frac{n-1}{n}\right)+\frac{2}{T}\left(\sum_{i=1}^{n-1}x_i^2-\sum_{i=1}^{n-1}x_ix_{i+1}\right). \qquad (30)$$

Воспользовавшись средним значением одного отсчета и (10), найдем среднее $\langle\hat{D}\rangle$

$$\left\langle\hat{D}\right\rangle=\frac{(n-1)}{n}D. \qquad (31)$$

Как видно из (31), оценка коэффициента диффузии является смещенной, однако, нормируя ее на среднее значение, получим несмещенную оценку

$$\hat{D}_0=\frac{n}{n-1}\hat{D}, \qquad (32)$$

где

$$\left\langle\hat{D}_0\right\rangle=D. \qquad (33)$$

Вычислим дисперсию нормированной оценки

$$Var\left[\hat{D}_0\right]=\left(\frac{n}{n-1}\right)^2Var\left[\hat{D}\right]. \qquad (34)$$

Воспользовавшись выражением для среднего значения i-го отсчета, а также (10), (29) и тем, что

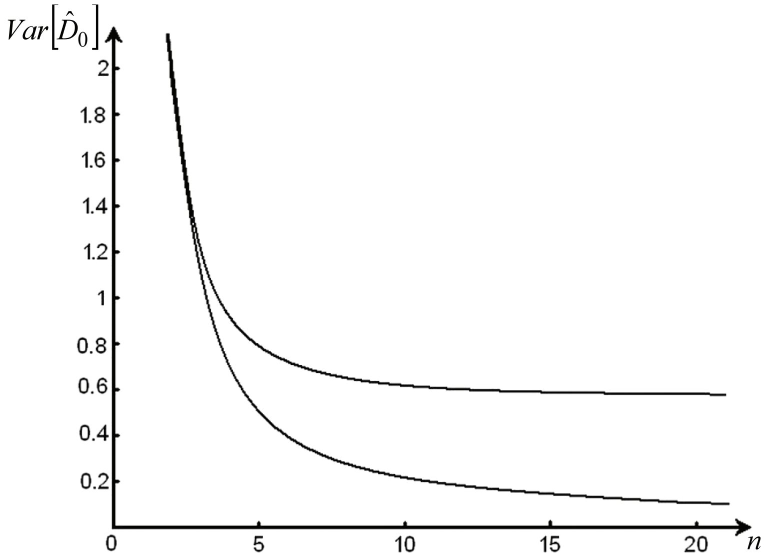

Рис. 3. Зависимость дисперсии оценки коэффициента диффузии, полученной методом наибольшего правдоподобия, от количества отсчетов n – нижняя линия. Для сравнения приведена оценка, полученная методом наименьших квадратов – верхняя линия

$$x_i = ct_i + \sqrt{D}W(t_i) = \frac{cT}{n}i + \sqrt{D}W\left(\frac{Ti}{n}\right), \quad (35)$$

найдем $\langle \hat{D}^2 \rangle$

$$\left\langle \hat{D}^2 \right\rangle = D^2 \frac{n^2 - 1}{n^2}. \quad (36)$$

Теперь можно подсчитать дисперсию ненормированной оценки

$$Var\left[\hat{D}\right] = 2D^2 \frac{n-1}{n^2}. \quad (37)$$

Используя (37), из (34) получим

$$Var\left[\hat{D}_0\right] = \frac{2}{n-1} D^2. \quad (38)$$

Как видно из рис. 3, дисперсия оценки коэффициента диффузии, полученная методом наибольшего правдоподобия, меньше, чем дисперсия оценки, полученной методом наименьших квадратов, а значит, оценка коэффициента диффузии методом наибольшего правдоподобия является более эффективной.

Заключение

В статье обсуждены два традиционных способа оценки параметров случайных процессов: метод наименьших квадратов и метод наибольшего правдоподобия. На примере рассматриваемого винеровского случайного процесса с равномерным сносом найдены оценки коэффициента диффузии и проверено соответствие полученной оценки критериям качества, а также рассмотрена ее асимптотика. В рассматриваемом случае оценка, полученная методом наибольшего правдоподобия, оказалась более эффективной, чем результат, полученный методом наименьших квадратов.

Наиболее актуальным неразрешенным вопросом остается вопрос о построении новых, более эффективных методов оценок параметров случайных процессов. Поскольку неизвестным остается тот факт, возможно ли построение более эффективных, чем полученные в данной статье, оценок другими методами, в дальнейшем планируется изучение различных методов оценивания параметров и построение универсальной формы оценки параметров случайных процессов.

Список литературы

1. Крамер Г. Математические методы статистики: Пер. с англ. М.: Мир, 1975. 645 с.

2. Ширяев А.Н. Основы стохастической финансовой математики. Т.1. Факты. Модели. М.: Фазис, 1998. 220 с.

3. Линник Ю.В. Метод наименьших квадратов и основы теории обработки наблюдений. М.: Физматгиз, 1962. 350 с.

4. Тихонов В.И. Статистическая радиотехника. М.: Радио и связь, 1982. 624 с.

5. Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. М.: Наука, 1974. 696 с.

DIFFUSION COEFFICIENT ESTIMATION
OF A WIENER RANDOM PROCESS WITH A UNIFORM DRIFT

A.I. Saichev, V.A. Filimonov, M.V. Tarakanova

Maximum likelihood and least squares methods are proposed for diffusion coefficient estimation of a random Wiener process with a uniform drift. Unbiasedness has been shown to exist and variances of obtained estimations have been found. The comparison of the estimations obtained has also been presented.

Keywords: statistical radiophysics, diffusion coefficient, statistical parameter estimation.