EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 31. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1982, NR. 3

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 31
ФИЗИКА * МАТЕМАТИКА. 1982, № 3

УДК 519.246 : 517.956.34

T. TOBIAS

IDENTIFICATION OF THE DIFFUSION PROCESS AND INVERSE PARABOLIC PROBLEMS

(Presented by N. Alumäe)

Let x_t be a d-dimensional diffusion process controlled by the equation $dx_t=b(x,\theta)\,dt+\sigma(x,\theta)\,dw_t$, $x(0)=x_0$. The coefficients of the equation are dependent on the unknown parameter θ. Usually in estimation problems it is supposed that the trajectory x_s, $0<s\leqslant T$, is observable at every moment s and then the estimate $\hat{\theta}$ can depend on the whole trajectory x_s, $0<s\leqslant T$.

In this work we shall consider a different model. Various trajectories of the process are observable only at separate random moments connected with the exit time of the process from the given region. For example, the observations can be the time intervals during which the various trajectories of the process have reached the boundary. Inside the region the trajectories are not observable. The observations on the boundary can be used for approximating certain functionals depending on the trajectories of the process and thus it is possible to estimate θ itself.

Such a problem was set up at [1] where some examples were given, permitting to derive analytical solutions.

In this paper we shall derive some approximate methods to estimate θ in more general models.

1. Let x_t be the homogeneous diffusion process with strictly elliptic infinitesimal operator $L=\sum\limits_{i,j=1}^{d} a_{ij}(x,\theta)\dfrac{\partial^2}{\partial x_i \partial x_j}+\sum\limits_{i=1}^{d} b_i(x,\theta)\dfrac{\partial}{\partial x_i}$ with smooth coefficients $a_{ij}(x,\theta)$ and $b_i(x,\theta)$. Let $D\subset R^d$ be a bounded set with smooth boundary Γ and let $\gamma\subset\Gamma$. Let $u_\gamma(t,x)$ be the probability that $x(s)$, $0<s\leqslant t$, $x(0)=x$ has reached the boundary γ during the time t. It is known [2] that $u=u_\gamma(t,x)$ is the solution of the equation

$$\frac{\partial u}{\partial t}=Lu, \tag{1}$$

$$u(0,x)=0,\quad u(t,x)=1,\quad x\in\gamma,\quad u(t,x)=0,\quad x\in\Gamma\backslash\gamma.$$

Suppose that it is possible to estimate the function $u(t,x)$. For example, suppose that over trajectory we can only determine if it has reached the boundary γ during the time t or not. If from N trajectories $k(t,x)$ has reached γ during t, then $u_N(t,x)=k(t,x)/N\approx u(t,x)$ and $\lim u_N(t,x)=u(t,x)$ with probability 1 for $x\in D$. Now the estimating problem can be stated as follows: given the estimate of the solution of the equation (1), find the estimates for the coefficients $a_{ij}(x)$ and $b_i(x)$ in nonparametric case or find the estimate for the parameter θ in parametric case.

The given model corresponds to the situation when the process is observable only on the part γ of the boundary Γ.

In such a setting the problem is closely connected with the inverse problems in differential equations [3, 4]. These inverse problems we shall treat as certain deterministic control problems. Such a point of view is well known. In the control theory for partial differential equations this method was used mainly in linear cases where the control (estimated parameter) contained in the right-hand side of the equation or in the initial (boundary) conditions of the equation [5]. The problem with the control in the coefficients is more complicated [5, 6] and the approximative methods are considered mainly in the case of equations with constant coefficients [7].

We shall learn the parametric case and in order to simplify the notations, only one-dimensional case is treated. We suppose that $a(x,\theta)=$ $=\sum_{i=1}^{n}\theta_i a_i(x)$ and $b(x,\theta)\equiv 0$. All the results and methods can be generalized without changes to the case $x\in R^d$ for estimating the parametrized functions $a_{ij}(x,\theta)$ and $b_i(x,\theta)$.

2. Let $a_1(x)>0, \ldots, a_n(x)>0$ be linearly independent functions on the interval $x\in[0,1]$. Denote $a'=(a_1, \ldots, a_n)$, $\theta'=(\theta_1, \ldots, \theta_n)$ and $(\theta,a)=\sum_{i=1}^{n}\theta_i a_i(x)$. Let us consider the equation

$$u_t-(\theta,a)\,u_{xx}=0, \tag{2}$$

$$u(0,x)=0, \quad u(t,0)=1, \quad u(t,1)=0. \tag{3}$$

Suppose that $\theta\in M$, where $M\subset R^n$ is a bounded closed set. Suppose that $a_i(x)$, $x\in[0,1]$ satisfy the Hölder condition with degree α, i.e. $a_i(x)\in C_\alpha$. Let $(\theta,a)\geqslant c>0$, $\theta\in M$, $x\in[0,1]$. It is known [8] that the equation (2), (3) has the unique solution and inside the region $D_T=$ $=[0,T]\times[0,1]$ $u\in C_{2+\alpha}$. If $\theta_m\to\theta$ then $u_m=u_m(t,x;\theta_m)$, $(u_m)_t$, $(u_m)_x$ and $(u_m)_{xx}$ converge to $u=u(t,x;\theta)$, u_t, u_x and u_{xx} uniformly on every open subregion of D_T.

Let $\bar{u}=u(t,\bar{x};\theta)=z_N(t)+\varepsilon_N(t)$, where $u(t,\bar{x};\theta)$ is the probability that trajectory of the diffusion process x_s, $0<s\leqslant t$, $x(0)=\bar{x}$, has reached the boundary $x=0$ during the time t and let $z_N(t)$ be the estimate of this probability obtained from N observations. If $k(t,x)$ from N trajectories has reached the boundary $x=0$ during t then $z_N(t)=k(t,\bar{x})/N$ and the error $\varepsilon_N(t)$ has asymptotically normal distribution with zero mean.

Suppose that $z_N(t)=z(t)$ is fixed. Let

$$J(\theta)=\int_0^T[u(t,\bar{x};\theta)-z(t)]^2dt.$$

The problem is to find $\theta\in M$ so that $J(\theta)\to\min$. Evidently, there exists $\theta^*\in M$ such that $\min_{\theta\in M}J(\theta)=J(\theta^*)$.

Remark 1. If M is unbounded, then θ^* need not exist. For example, let $u_t-\theta u_{xx}=0$ with conditions (3) and let $z(t)\equiv 1$. If θ_n is the minimizing sequence, then $\theta_n\to\infty$.

3. To solve the problem (to find the optimal θ^*) it is necessary to use some approximative numerical method. Corresponding effective procedures, as a rule, use the gradient of the function $J(\theta)$. We shall derive the exact expression for the $(\operatorname{grad}J(\theta),\Delta\theta)=\sum_{i=1}^{n}\frac{\partial J(\theta)}{\partial\theta_i}\Delta\theta_i$.

Lemma 1. *Denote* $v_i = v_i(t, x; \theta) = \frac{\partial u(t, x; \theta)}{\partial \theta_i}$. *Then* $v_i(t, x; \theta)$ *exists and satisfies the equation:*

$$(v_i)_t - (\theta, a)(v_i)_{xx} = a_i u_{xx},$$
$$v_i(0, x; \theta) = 0; \quad v_i(t, 0; \theta) = v_i(t, 1; \theta) = 0, \quad i = 1, \ldots, n. \tag{4}$$

Denote by $u^\Delta(t, x; \theta)$ the solution of the equation $u^\Delta_t - (\theta + \Delta\theta, a) u^\Delta_{xx} = 0$ with the conditions (3). Let $v^\Delta_i = (u^\Delta - u)/\Delta\theta_i$, then

$$(v^\Delta_i)_t - (\theta, a)(v^\Delta_i)_{xx} = \frac{(\Delta\theta, a)}{\Delta\theta_i} u^\Delta_{xx},$$
$$v^\Delta_i(0, x) = 0, \quad v^\Delta_i(t, 0) = v^\Delta_i(t, 1) = 0. \tag{5}$$

If $\Delta\theta \to 0$, then u^Δ_{xx} converges uniformly to u_{xx} in every open subregion of $[0, 1]$ and thus v^Δ_i with corresponding derivatives converges uniformly [8]. This proves the lemma.

It is easy to see that $\operatorname{grad} J(\theta) = 2\int\limits_0^T [u(t, \bar{x}; \theta) - z(t)] v(t, \bar{x}; \theta) dt$, where $v'(t, x; \theta) = (v_1, \ldots, v_n)$.
Let us transform this expression.

Let $p = p(t, x; \theta)$ be the solution of the conjugate equation:

$$-\frac{\partial p}{\partial t} - \frac{\partial^2}{\partial x^2}[(\theta, a)p] = 2[u(t, x; \theta) - z(t)]\delta(x - \bar{x}),$$
$$p(T, x; \theta) = 0, \quad p(t, 0; \theta) = p(t, 1; \theta) = 0, \tag{6}$$

where $\delta(x - \bar{x})$ is δ-function. By multiplying the equation (6) with the vector $v(t, x; \theta)$ and integrating each component over the region D_T, we get (using the equation (4)):

$$\operatorname{grad} J(\theta) = \int\limits_0^T \int\limits_0^1 [v_t - (\theta, a) v_{xx}] p \, dx \, dt =$$

$$= \int\limits_0^T \int\limits_0^1 a(x) u_{xx}(t, x; \theta) p(t, x; \theta) \, dx \, dt.$$

Finally:

$$(\operatorname{grad} J(\theta), \Delta\theta) = \sum_{i=1}^n \int\limits_0^T \int\limits_0^1 p(t, x; \theta) u_{xx}(t, x; \theta) a_i(x) \, dx \, dt \, \Delta\theta_i, \tag{7}$$

where $u(t, x; \theta)$ and $p(t, x; \theta)$ are the solutions of the equations (2), (3) and (6).

By means of the expression (7) we can construct various numerical procedures to minimize the function $J(\theta)$.

Remark 2. Such a method for calculating the gradient by using the conjugate equations is known and was used, for example, in [6].

4. Let us consider the identifiability of the parameter θ. Suppose that the values of the function $u = u(t, x;\ \theta)$ are known at fixed points $(t_1, \bar{x}), \ldots, (t_n, \bar{x})$, i.e. the values $u_1 = u_1(t_1, \bar{x};\ \theta), \ldots, u_n = u_n(t_n, \bar{x};\ \theta)$ are known. The corresponding Jacobian is

$$D(u_1, \ldots, u_n) = \begin{vmatrix} v_1(t_1, \bar{x}; \theta) & \ldots & v_n(t_1, \bar{x}; \theta) \\ \ldots & \ldots & \ldots \\ v_1(t_n, \bar{x}; \theta) & \ldots & v_n(t_n, \bar{x}; \theta) \end{vmatrix}.$$

If $D(u_1, \ldots, u_n) \neq 0$, then θ is identifiable.

The functions $v_1, \ldots, v_n$ are independent in the whole region D_T because $a_1(x), \ldots, a_n(x)$ are independent. But in certain subregion they may be dependent and therefore, in general, the values $u_1, \ldots, u_n$ do not determine the parameter θ.

Consider the special case $n=1$. Let

$$\begin{aligned} &u_t-(a_0(x)+\theta a_1(x))u_{xx}=0, \\ &u(0,x)=0,\ u(t,0)=1,\ u(t,1)=0, \end{aligned} \tag{8}$$

where $a_0(x)>0$ and $a_1(x)>0$ are known functions and θ is an unknown parameter. $a_0(x)$ can be interpreted as the approximate value of the unknown function $a(x)$ and $a_1(x)$ is the correction.

Lemma 2. *For every $t\in(0, T)$ and $\bar{x}\in(0, 1)$ the value $u=u(t,\bar{x};\theta)$ determines the parameter θ uniquely.*

The function $v(t,x;\ \theta)=\dfrac{\partial u}{\partial\theta}$ satisfies the equation

$$\begin{aligned} &v_t-(a_0(x)+\theta a_1(x))v_{xx}=a_1(x)u_{xx}, \\ &v(0,x)=0,\ v(t,0)=v(t,1)=0. \end{aligned}$$

Denote by $G(x,\xi;\ t)$ the Green's function of the equation (8); then

$$v(t,x;\ \theta)=\int_0^t\int_0^1 G(x,\xi;\ t-\tau)a_1(\xi)u_{xx}(\tau,\xi;\ \theta)\,d\xi\,d\tau.$$

As $u(t,x;\ \theta)$ is the probability of reaching the boundary $x=0$ during the time t, then $u_t\geqslant0$ and, therefore, $u_{xx}\geqslant0$ (evidently $u_{xx}\not\equiv0$ on the whole interval $(0, 1)$). So, $v(t,x;\ \theta)>0$ and this proves the lemma.

The derivative of the function $J(\theta)$ can be computed as $\dfrac{d}{d\theta}J(\theta)=$ $=\int_0^T\int_0^1 u_{xx}(t,x;\ \theta)p(t,x;\ \theta)a_1(x)\,dx\,dt$, where $p(t,x;\ \theta)$ satisfies the equation

$$\begin{aligned} &-\frac{\partial p(t,x;\theta)}{\partial t}-\frac{\partial^2}{\partial x^2}[a_0(x)+\theta a_1(x)p(t,x;\theta)]= \\ &=2[u(t,x;\theta)-z(t)]\delta(x-\bar{x}), \\ &p(T,x)=0,\ p(t,0)=p(t,1)=0. \end{aligned}$$

5. Instead of the exact value $u(t,\bar{x};\ \theta)$ we have only the approximate value $z_N(t)$. Suppose that $z_N(t)\to u(t,\bar{x};\ \theta^*)$ with probability 1, where θ^* is the true value of the parameter θ.

Let $J_N(\theta)=\int_0^T[u(t,\bar{x};\ \theta)-z_N(t)]^2dt$ and $\min\limits_{\theta\in M}J_N(\theta)=J(\theta_N)$.

Lemma 3. *Let $\theta^*\in M$. Suppose that if $u(t,\bar{x};\ \theta_1)=u(t,\bar{x};\ \theta_2)$, $0\leqslant t\leqslant T$, then $\theta_1=\theta_2$. There exists subsequence $\{\theta_{N_t}\}$ such that* $\lim\theta_{N_t}=\theta^*$ *with probability* 1.

As $\theta_{N_t}\in M$, we can find $\{\theta_{N_t}\}$ such that $\theta_{N_t}\to\tilde{\theta}$. By continuity $u(t,\bar{x};\ \theta_{N_t})\to u(t,x;\ \tilde{\theta})$. As $\lim z_{N_t}(t)=u(t,\bar{x};\ \theta^*)$ with probability 1 and $J_{N_t}(\theta^*)\to0$, then $J_{N_t}(\theta_{N_t})\to0$. But then $J_{N_t}(\tilde{\theta})\to0$ with probability 1 and $u(t,\bar{x};\ \theta^*)=u(t,\bar{x};\ \tilde{\theta})$. By assumption we conclude that $\theta^*=\tilde{\theta}$.

6. We have derived the exact value for the gradient of the function $J(\theta)$. In order to use this expression it is necessary to solve at each

step two partial differential equations. Now we shall approximate the initial equation by Galerkin's method and then we shall derive the expression for the corresponding gradient. At every step its application will insist on solving two ordinary differential equations.

Let $U(t,x;\ \theta)=u(t,x;\ \theta)+(x-1)$, then instead of (2) and (3) we have

$$U_t-(\theta,a)U_{xx}=0, \tag{9}$$

$$U(0,x;\ \theta)=x-1, \quad U(t,0;\ \theta)=U(t,1;\ \theta)=0.$$

Let $\{f_n(x)\}$, $x\in[0,1]$ be the complete set of basis functions, $f_n(0)=$ $=f_n(1)=0$. We search the approximate solution in the form $U(t,x;\ \theta)\approx$ $\approx u_m(t,x;\ \theta)=\sum_{i=1}^{m}\varphi_i(t)f_i(x)$. Putting the expression $u_m(t,x;\ \theta)$ into the equation (9), multiplying the equation (9) by $f_k(x)$ and integrating we get the following equation for the function $\varphi_k(t)$:

$$\frac{d\varphi_k(t)}{dt}=\sum_{i=1}^{n}\sum_{j=1}^{m}\theta_i a_{ij}^{k}\varphi_j(t), \tag{10}$$

$$\varphi_k(0)=c_k, \quad k=1,\ \dots,\ m.$$

Here $a_{ij}^{k}=\int_0^1 a_i(x)f_j''(x)f_k(x)\,dx$ and $x-1=\sum_{k=1}^{\infty}c_kf_k(x)$. It is known that $u_m\to U$ uniformly inside the region D_T [9]. Let $A_k=(a_{ij}^{k})$, $\varphi'=(\varphi_1(t),\ \dots$ $\dots,\ \varphi_m(t))$, $c'=(c_1,\ \dots,\ c_m)$. Denote $A(\theta)=\begin{pmatrix}\theta'A_1\\ \cdots\\ \theta'A_m\end{pmatrix}$. Then the system (12) can be transformed to the form

$$\frac{d\varphi}{dt}=A(\theta)\varphi, \tag{11}$$

$$\varphi(0)=c.$$

The function $J_m(\theta)$ can be expressed as

$$J_m(\theta)=\int_0^T[(d,\varphi)-\bar z(t)]^2\,dt,$$

where $d'=(f_1(\bar x),\ \dots,\ f_m(\bar x))$, $\bar z(t)=z(t)+(\bar x-1)$. Denote $v_k(t)=$ $=\left(\frac{d\varphi_k}{d\theta},\ \Delta\theta\right)$, $v'=(v_1(t),\ \dots,\ v_m(t))$. It is easy to see that

$$\frac{dv}{dt}=Av+B\Delta\theta, \quad v(0)=0, \tag{12}$$

where $B=\begin{pmatrix}\varphi'A'_1\\ \cdots\\ \varphi'A'_m\end{pmatrix}$ and $'$ denotes the transposition. Let us calculate the gradient of the function $J_m(\theta)$

$$(\operatorname{grad}J_m(\theta),\ \Delta\theta)=2\int_0^T[(d,\varphi)-\bar z(t)]\left(d'\frac{d\varphi}{d\theta}\Delta\theta\right)dt=2\int_0^T[(d,\varphi)-$$ $$-\bar z(t)](d,v)\,dt.$$

Introduce the conjugate system

$$-\frac{dp}{dt}-A'p=2[(d,\varphi)-\bar z(t)]d, \quad p(T)=0, \tag{13}$$

where $A' = (A'_1\theta, \ldots, A'_m\theta)$ and $p' = (p_1(t), \ldots, p_m(t))$. Multiplying the equation (13) by $v(t)$ and integrating over $[0, T]$ we get $2\int_0^T ([(d,\varphi) -$

$$-\bar{z}(t)]d,\ v)\,dt = \int_0^T \left(\left[-\frac{dp}{dt} - A'p\right],\ v\right)dt = \int_0^T \left(p,\ \frac{dv}{dt} - Av\right)dt =$$

$$= \int_0^T (p, B\Delta\theta)\,dt = \int_0^T (B'p, \Delta\theta)\,dt.$$

Finally we get

$$(\operatorname{grad} J_m(\theta), \Delta\theta) = \int_0^T (B'p(t), \Delta\theta)\,dt, \tag{14}$$

where $B' = (A_1\varphi, \ldots, A_m\varphi)$. Formula (16) is the basic formula in various numerical procedures for minimizing the function $J(\theta)$.

7. Let $\min\limits_{\theta\in M} J_m(\theta) = J_m(\theta_m)$.

Lemma 4. *If* $\theta^* \in M$, *then* $\lim J_m(\theta_m) = J(\theta^*)$.

It is clear that $J_m(\theta_m) \leqslant J_m(\theta^*) \to J(\theta^*)$, therefore $\overline{\lim}\, J_m(\theta_m) \leqslant J(\theta^*)$. Let $\{\theta_{m_i}\}$ be the converging subsequence, $\theta_{m_i} \to \tilde{\theta}$. By Fatou lemma: $\underline{\lim}\, J_{m_i}(\theta_{m_i}) \geqslant J(\tilde{\theta})$, therefore, $\overline{\lim}\, J_{m_i}(\theta_{m_i}) \leqslant J(\theta^*) \leqslant J(\tilde{\theta}) \leqslant \underline{\lim}\, J_{m_i}(\theta_{m_i})$ and $J(\theta^*) = J(\tilde{\theta})$.

REFERENCES

1. Тобиас Т., Изв. АН ЭССР, Физ. Матем., **29**, № 4, 432—435 (1980).
2. Тихонов В. И., Выбросы случайных процессов, М., «Наука», 1970.
3. Лаврентьев М. М., Романов В. Г., Шишатский С. П., Некорректные задачи математической физики и анализа, М., «Наука», 1980.
4. Успенский А. Б., В кн.: Математические методы планирования эксперимента, Новосибирск, «Наука», 1981, с. 199—242.
5. Лионс Ж., Оптимальное управление системами, описываемыми уравнениями с частными производными, М., «Мир», 1972.
6. Lions, J. L., In: Lecture Notes in Control and Information Sciences, **1**, Springer, Berlin, 1978, p. 11—41.
7. Polis, M. P., Goodson, R. E., Wozny, M. J., Automatica, **9**, № 1, 53—64 (1973).
8. Фридман А., Уравнения с частными производными параболического типа, М., «Мир», 1968.
9. Ильин А. М., Калашников А. С., Олейник О. А., Успехи мат. наук, **18**, вып. 3, 3—146 (1962).

Academy of Sciences of the Estonian SSR,
Institute of Cybernetics

Received
Dec. 30, 1981

T. TOBIAS

DIFUSIOONIPROTSESSI IDENTIFITSEERIMINE JA PARABOOLSET TÜÜPI VÕRRANDITE PÖÖRDÜLESANDED

On uuritud difusiooniprotsessi kordajate hindamist, kui vaatlustulemusteks on protsessi erinevate trajektooride antud piirkonnast väljumise momendid. On näidatud, et ülesanne on ekvivalentne teatud pöördülesandega paraboolset tüüpi võrrandite teoorias, kus lahendi väärtuste põhjal tuleb leida võrrandi kordajad. Seda pöördülesannet on vaadeldud kui teatud minimeerimisülesannet. On esitatud avaldis sihifunktsiooni gradiendi arvutamiseks ning vaadeldud Galjorkini meetodil põhinevat diskretiseerimismeetodit.

Т. ТОБИАС

ИДЕНТИФИКАЦИЯ ДИФФУЗИОННОГО ПРОЦЕССА И ОБРАТНЫЕ ПАРАБОЛИЧЕСКИЕ ЗАДАЧИ

Рассматривается задача оценивания параметрически заданных коэффициентов диффузионного процесса по наблюдениям на границе заданной области. Показывается, что задача эквивалентна обратным задачам в теории параболических уравнений, коэффициенты в которых требуется определить по значениям решения в определенных точках. Эта задача идентификации рассматривается как задача минимизации. Дается формула для вычисления градиента целевой функции. Рассматривается также метод дискретизации, основанный на методе Галеркина.