III ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК 1
LIETUVOS MATEMATIKOS RINKINYS

1963

К ОПТИМАЛЬНОМУ РЕГУЛИРОВАНИЮ ВИНЕРОВСКОГО ПРОЦЕССА

В. А. КОЛЕМАЕВ

1. Пусть $\xi(t)$ — случайный процесс, удовлетворяющий стохастическому дифференциальному уравнению

$$d\xi=\theta\,dt+\sigma\,d\eta,\quad \xi(0)=0, \tag{1}$$

где $\eta(t)$ — винеровский, т. е. гауссовский процесс с независимыми приращениями, $\eta(0)=0$, $\mathbf{M}\Delta\eta=0$, $\mathbf{D}\Delta\eta=\Delta t$.

В настоящей заметке рассматриваются некоторые задачи регулирования случайного процесса $\xi(t)$ при различных предположениях относительно параметра θ.

В дальнейшем под регулированием, осуществляемом на основе наблюдаемых данных, понимается мгновенный сдвиг процесса в первоначальное положение (подналадка), после чего на новом этапе сам процесс и наблюдение за ним возобновляются заново и происходят по тем же самым законам и независимо от прежних этапов.

Требуется производить подналадки в такие моменты времени, чтобы при некоторых условиях управляемый процесс как можно дольше оставался в заданном интервале $(-a,\ a)$.

Пусть X — пространство непрерывных $[0,\ \infty)$ функций $x=\{x(t),\ t\geqslant 0\}$ и $\mathfrak{B}_t$ — σ — алгебра множеств, порожденная функциями $x^t=\{x(s),\ s\leqslant t\}$. Управляющее правило δ задаётся системой функционалов $\{\Phi(t,\ x)\}$, $\mathbf{B}_t$ — измеримых при каждом t, при фиксированном x являющихся функциями распределения по t и удовлетворяющих естественным условиям согласованности (см., напр. [5]). Тем самым при каждом x определяется случайная величина $\tau(x)$, такая, что

$$\mathbf{P}\{\tau(x)\leqslant t\}=\Phi(t,\ x). \tag{2}$$

Пусть далее

$$V(z)=\begin{cases}1 & |z|<a,\\ 0 & |z|<a.^{*}\end{cases}$$

Тогда для каждого управляющего правила δ и заданного θ определено среднее время пребывания процесса $\xi(t)$ внутри интервала $(-a,\ a)$ до первой подналадки:

$$r(\delta,\ \Theta)=\mathbf{M}_{\Theta}\left\{\int\limits_0^{\tau} V[\xi(t)]\,dt/\xi(0)=0\right\}. \tag{3}$$

---

* Приводимые ниже дифференциальные уравнения остаются в силе и для произвольной ограниченной по модулю кусочнонепрерывной функции $V(x)$.

Наша задача состоит в том, чтобы среди всех правил, подчиняющихся условию

$$\mathbf{M}\tau = T, \tag{4}$$

найти ,,оптимальное" на одном этапе управляющее правило δ^*, т. е. правило, для которого

$$\sup_{\{\delta:\mathbf{M}\tau=T\}} r(\delta,\ \Theta) = r(\delta^*,\ \Theta) = \mathfrak{I}(T,\ \Theta). \tag{5}$$

Пусть $\nu(\delta,\ s)$ — время, проведенное регулируемым согласно правилу δ процессом на отрезке времени $[0,\ s)$ в интервале $(-a,\ a)$, и $\nu_i(\delta)$ — время, проведенное регулируемым согласно правилу δ процессом на отрезке времени $\left[\sum\limits_{k=1}^{i-1}\tau_k,\ \sum\limits_{k=1}^{i}\tau_k\right)$, где τ_k — промежуток времени между $(k-1)$ и k-ой подналадками.

Заметим, что правило δ^* будет также оптимальным и при решении следующей многоэтапной задачи:

$$\sup_{\{\delta:M\tau=T\}} \lim_{s\to\infty} \frac{\mathbf{M}\nu(\delta,\ s)}{s} = \lim_{s\to\infty} \frac{\mathbf{M}\nu(\delta^*,\ s)}{s}, \tag{6}$$

так как согласно [3]

$$\lim_{s\to\infty} \frac{\mathbf{M}\nu(\delta,\ s)}{s} = \lim_{s\to\infty} \frac{\mathbf{M}\left[\nu_1(\delta)+\ldots+\nu_{k-1}(\delta)+\nu_k'(\delta)\right]}{\tau_1(\delta)+\ldots+\tau_{k-1}(\delta)+\tau_k'(\delta)} = \frac{r(\Theta,\ \delta)}{\mathbf{M}\tau},$$

где штрихи означают, что до момента s k-ая подналадка еще не произошла.

2. Найдем теперь оптимальное правило в предположении, что θ известно. Используя метод множителей Лагранжа, можно свести задачу (5) нахождения условного экстремума к задаче отыскания правила δ^*, для которого

$$\rho(\delta^*,\ \theta) = \sup_{\{\delta\}} \rho(\delta,\ \theta), \tag{7}$$

где

$$\rho(\delta,\ \theta) = \mathbf{M}_\Theta\left\{\int\limits_0^\tau V[\xi(t)]\,dt - c\tau/\xi(0)=0\right\}.$$

Отметим, что случаи $c\leqslant 0$, $c\geqslant 1$ тривиальны, поэтому далее будем считать $0<c<1$.

Пусть

$$\rho(z) = \sup_{\{\delta\}} \mathbf{M}_\Theta\left\{\int\limits_0^\tau V[\xi(t)]\,dt - c\tau/\xi(0)=z\right\}. \tag{8}$$

Если рассматривать только регулярные правила (см. [1]) и только те из них, для которых функция $\rho\ (z)$ дважды дифференцируема, то можно показать, что $\rho\ (z)$ является непрерывно-дифференцированным решением уравнения:

$$\frac{\sigma^2}{2}\,\rho'' + \theta\rho' + V(z) - c = 0, \tag{9}$$

удовлетворяющим начальным условиям

$$\rho(\beta) = \rho(\alpha) = 0,\ \beta < -a < a < \alpha, \tag{10}$$

где α, β определяются из условия

$$\rho'(\beta)=\rho'(\alpha)=0 \tag{10'}$$

и $\rho(z)=0$ для $z\overline{\in}(\beta,\ \alpha)$.

В результате решения уравнения (9) с условиями (10), (10′) найдём

$$\rho(0)=\frac{\sigma^2}{2\theta^2}\left[1-\frac{sh\left(\frac{2\theta(\alpha+\Delta-a)}{\sigma^2}\right)+\exp\left(\frac{2\theta\Delta}{\sigma^2}\right)\operatorname{sh}\left(\frac{2\theta a}{\sigma^2}\right)}{\operatorname{sh}\left(\frac{2\theta(\alpha+\Delta)}{\sigma^2}\right)}\right]+$$

$$+\frac{1}{\theta}\left[c\Delta+(a-c\alpha-c\Delta)\frac{\operatorname{ch}\left(\frac{2\theta(\alpha+\Delta)}{\sigma^2}\right)-\exp\left(-\frac{2\theta\Delta}{\sigma^2}\right)}{\operatorname{sh}\left(\frac{2\theta(\alpha+\Delta)}{\sigma^2}\right)}\right], \tag{11}$$

где

$$\Delta=\frac{\sigma^2}{4\theta}\ln\frac{c\operatorname{sh}2\theta\frac{a}{c\sigma^2}}{\operatorname{sh}2\theta\frac{a}{\sigma^2}}\quad \alpha=\frac{a}{c}-2\Delta,\ \ \beta=-\frac{a}{c}-2\Delta. \tag{12}$$

В выражении (11) коэффициент при $(-c)$ представляет собой согласно (7)

$$\mathbf{M}\tau=\frac{1}{\theta}\left[-\Delta+(\alpha+\Delta)\frac{\operatorname{ch}\frac{2\theta}{\sigma^2}(\alpha+\Delta)-e^{-\frac{2\theta\Delta}{\sigma^2}}}{\operatorname{sh}\frac{2\theta}{\sigma^2}(\alpha+\Delta)}\right]=T. \tag{13}$$

Решая совместно уравнения (12) и (13), найдем

$$\alpha=\alpha(T,\ \theta),\ \ \beta=\beta(T,\ \theta),\ \ \Delta=\Delta(T,\ \theta),\ \ c=c(T,\ \theta) \tag{14}$$

и

$$\mathfrak{I}(T,\ \theta)=\frac{\sigma^2}{2\theta^2}\left[1-\frac{\operatorname{sh}\frac{2\theta(\alpha+\Delta-\theta)}{\sigma^2}+\exp\left(\frac{2\theta\Delta}{\sigma^2}\right)\operatorname{sh}\left(\frac{2\theta a}{\sigma^2}\right)}{\operatorname{sh}\left(\frac{2\theta(\alpha+\Delta)}{\sigma^2}\right)}\right]+$$

$$+\frac{a}{\theta}\cdot\frac{\operatorname{ch}\frac{2\theta(\alpha+\Delta)}{\sigma^2}-\exp\left(-\frac{2\theta\Delta}{\sigma^2}\right)}{\operatorname{sh}\frac{2\theta(\alpha+\Delta)}{\sigma^2}}, \tag{15}$$

где α и Δ взяты из (14), при этом само оптимальное правило задается границами

$$\alpha=\alpha(T,\ \theta),\ \ \beta=\beta(T,\ \theta).$$

В случае $\theta=0$ можно найти явный вид для функций (14) и (15)

$$\alpha=-\beta=\sqrt{T},\ \ \Delta=0,\ \ c=\frac{a}{\sqrt{T}},$$

$$\mathfrak{I}(T)=2a\sqrt{T}-a^2.$$

Для сравнения укажем, что если $\theta=0$ и подналадки производятся независимо от хода процесса через промежутки времени, распределенные по показательному закону с параметром $\frac{1}{T}$, то получим

$$\mathfrak{I}(T)=T\left(1-e^{-a\sqrt{\frac{2}{T}}}\right).$$

3. Пусть теперь для параметра θ задано априорное распределение $\pi(d\theta)$, сосредоточенное на отрезке $[\theta_0,\ \theta_1]$, $-\infty<\theta_0<\theta_1<\infty$.

В этом случае естественно рассматривать осреднённые величины „доходов“

$$R(\delta, \pi)=\int_{\Theta_0}^{\Theta_1} r(\theta, \delta)\,\pi(d\theta), \quad P(\delta, \pi)=\int_{\Theta_0}^{\Theta_1} \rho(\theta, \delta)\,\pi(d\theta).$$

Пусть

$$\pi_{z,t}(d\theta)=\frac{\exp\left\{\theta z-\frac{\theta^2}{2}t\right\}\pi(d\theta)}{\int_{\Theta_0}^{\Theta_1}\exp\left\{\theta z-\frac{\theta^2}{2}t\right\}\pi(d\theta)} \tag{16}$$

апостериорное распределение для параметра θ, в предположении, что $\xi(t)=z$. Тогда, используя метод [1], можно показать, что апостериорный „доход“

$$\rho(t, z)=\max_{\{\delta\}}\int_{\Theta_0}^{\Theta_1}\mathbf{M}_\Theta\left\{\int_0^{\tau} V[\xi(t)]\,dt-c\tau/\xi(0)=z\right\}\pi(d\theta) \tag{17}$$

для $(t, z)\in G$ является ограниченным решением дифференциального уравнения

$$\frac{\partial\rho}{\partial t}+\frac{\sigma^2}{2}\frac{\partial^2\rho}{\partial z^2}+g(z, t)\frac{\partial\rho}{\partial z}+V(z)-c=0), \tag{18}$$

$$g(z, t)=\int_{\Theta_1}^{\Theta_0}\theta\pi_{z,t}(d\theta)$$

при краевых условиях

$$\rho[t, \gamma_1(t)]=\rho[t, \gamma_2(t)]=0, \quad \gamma_1(t)<-a<a<\gamma_2(t),$$

где $\gamma_i(t)$ находятся из условия

$$\left.\frac{\partial\rho}{\partial z}\right|\gamma_1=\left.\frac{\partial\rho}{\partial z}\right|\gamma_2=0$$

и $\rho(t, z)=0$ для $(t, z)\overline{\in}G$, где область G заключена внутри $\gamma_1(t)$, $\gamma_2(t)$.

Если априорное распределение сосредоточено в двух симметричных точках, то согласно (16) апостериорное распределение не будет зависеть от t, поэтому оптимальными границами будут прямые.

Что касается конкретного вида кривых (γ_1, γ_2) в общем случае, то можно лишь утверждать, что они имеют горизонтальные асимптоты. В самом деле, согласно (16) апостериорное распределение для всех $|z|<K<\infty$ при $t\to\infty$ стягивается в одну или две (симметричных) точки, а для априорных распределений, сосредоточенных в одной или двух симметричных точках, оптимальными границами являются прямые.

Трудность нахождения оптимальной границы в общем случае заставляет пойти по пути нахождения наименее благоприятных априорных распределений, т. е. таких распределений π^*, что

$$\sup_{\{\delta\}}\mathbf{P}(\delta, \pi^*)=\inf_{\{\pi\}}\sup_{\{\delta\}}\mathbf{P}(\delta, \pi).$$

На этот счет справедлива следующая

Теорема 1. *Априорные распределения*

1) $\pi_1^*(\Theta^*)=1$, *если* $-\theta_0\geqslant 0$ (обозначено $\theta^*=\theta_1$);

2) $\pi_2^*(\theta)=\begin{cases}\frac{1}{2} & \theta=\theta_0\\ \frac{1}{2} & \theta=\theta_1,\ \text{если}\ -\theta_0=\theta_1\end{cases}$

являются наименее благоприятными.

Доказательство. Можно показать, что „доход“ $\rho(\theta, \alpha, \beta)$ для прямолинейных границ $a < \alpha < \frac{a}{c}$, $\beta \leqslant -\alpha$ является при $\theta \geqslant 0$ монотонно-убывающим как функция θ, поэтому справедливы следующие неравенства:

$$\sup_{\{\delta\}} \mathbf{P}(\delta, \pi) \geqslant \sup_{\{\alpha, \beta\}} \int_{\Theta_0}^{\Theta'} \rho(\theta, \alpha, \beta)\, \pi(d\theta) \geqslant \int_{\Theta_0}^{\Theta_1} \rho(\theta, \alpha^*, \beta^*)\, \pi(d\theta) \geqslant$$
$$\geqslant \rho(\theta^*, \alpha^*, \beta^*) = \sup_{\{\delta\}} \mathbf{P}(\delta, \pi_1^*),$$

и так как π — любое, то первый случай доказан. Во втором случае имеют место следующие неравенства

$$\sup_{\{\delta\}} P(\delta, \pi) \geqslant \sup_{\{\alpha=-\beta\}} \int_{\Theta_0}^{\Theta_1} \rho(\theta, \alpha, -\alpha)\, \pi(d\theta) \geqslant \int_{\Theta_0}^{\Theta_1} \rho(\theta, \alpha^*, -\alpha^*)\, \pi(d\theta) \geqslant$$
$$\geqslant \rho(\theta, \alpha^*, -\alpha^*) = \sup_{\{\delta\}} P(\delta, \pi_2^*),$$

что и требовалось доказать.

4. Пусть теперь относительно параметра θ известно только, что $\theta \in [\theta_0, \theta_1]$. Тогда естественно решать минимаксную задачу

$$\sup_{\{\delta\}} \inf_{\theta \in [\theta_0, \theta_1]} \rho(\delta, \theta) = \inf_{\theta \in [\theta_0, \theta_1]} \rho(\delta^*, \theta).$$

Справедлива следующая

Теорема 2. *Правило δ^*, являющееся оптимальным для наименее благоприятного распределения π_1^* является одновременно минимаксным.*

Действительно для любого правила δ

$$\rho(\theta, \delta) \leqslant \rho[\theta, \alpha^*(\theta), \beta^*(\theta)],$$

поэтому

$$\inf_{\theta \in [\theta_0, \theta_1]} \rho(\theta, \delta) \leqslant \inf_{\theta \in [\theta_0, \theta_1]} \rho[\theta, \alpha^*(\theta), \beta^*(\theta)] = \rho(\theta^*, \alpha^*, \beta^*) = \sup_{\{\delta\}} \mathbf{P}(\delta, \pi_1^*),$$

т. к. это верно для любого правила δ, то теорема доказана. Отметим, что последняя теорема аналогична теореме (5.4.) Вальда в [4] стр. 317.

В терминах $r(\delta, \theta)$ правило δ^* также будет минимаксным в следующем смысле

$$\sup_{\{\delta: M_\tau \geqslant T\}} \inf_{\theta \in [\theta_0, \theta_1]} r(\delta, \theta) = \inf_{\theta \in [\theta_{01}, \theta_1]} r(\delta^*, \theta).$$

Из теоремы 2 и результатов п. 2 вытекает, что минимаксное правило состоит в наблюдении значений процесса $\xi(t)$ до первого достижения одной из двух прямолинейных границ $\alpha = \alpha^*$, $\beta = \beta^*$.

В заключение выражаю благодарность А. Н. Ширяеву за постановку данной задачи и ценные советы при ее решении.

Поступила в редакцию
29.XI.1962

ЛИТЕРАТУРА

1. В. С. Михалевич. Последовательные байесовские решения и оптимальные методы контроля, Теория вероят. и ее примен., **1**, 4 (1956), 437—465.
2. Р. З. Хасьминский. Распределение вероятностей для функционалов от траектории случайного процесса диффузионного типа, ДАН СССР, **104**, 1 (1955), 22—25.

3. Smith W. L. Renewal theory and its ramifications. Jornal of the Rogal Statistical Society, B, 20 (1958), N 2, 243–284.
4. А. Вальд. Последовательный анализ. М. физ.-матгиз, 1960.
5. А. Н. Ширяев. Задача скорейшего обнаружения нарушения стационарного режима. Канд. диссертация, МИАН, М., 1961.

VINERIO PROCESO OPTIMALAUS VALDYMO KLAUSIMU

V. A. KOLEMAJEVAS

(Reziumė)

Darbe nagrinėjamas Vinerio proceso $\xi(t)$, tenkinančio lygtį (1), optimalaus valdymo uždavinys, kai apie parametrą θ daromos įvairios prielaidos.

Valdymo tikslas—kuo „ilgesnį" laiką išlaikyti stebimą procesą duotame intervale.

§ 2 nagrinėjamas atvejis, kai parametras θ žinomas tiksliai; § 3 tiriamas Baieso sprendinys; § 4 nagrinėjamos minimaksinės valdymo taisyklės.

ON AN OPTIMAL CONTROL OF THE WIENER PROCESS

V. A. KOLEMAJEV

(Summary)

A problem of the optimal control of the Wiener process $\xi(t)$, which satisfies an equation (1), is considered, when various assumptions about parameter θ are realised.

An aim of the control is to keep our process a „maximal" time in the given limits.

Results of § 2 deals with the case, when parameter θ' is known exactly in § 3 it is considered a Bayes desision, in § 4 there are examined the minimax rules of the control.