III ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК 1
LIETUVOS MATEMATIKOS RINKINYS
1963

ОБ ОДНОЙ ЗАДАЧЕ ОПТИМАЛЬНОГО РЕГУЛИРОВАНИЯ ВИНЕРОВСКОГО ПРОЦЕССА

В. А. КОЛЕМАЕВ

На интервале $(-a, a)$ наблюдается процесс $\xi=\{\xi_t\}$, $\xi_0=0$, принадлежащий семейству $X(\vartheta)$ диффузионных процессов с единичными локальными диффузиями и локальными сносами $\{\vartheta(t, x_t),\ t\geqslant 0,\ -a<x_t<a\}$, где функции $\vartheta(t, x_t)$ принимают значения из некоторого заданного множества Θ.

Пусть ν_a — момент первого выхода траектории x_t процесса $\xi=\{\xi_t\}$ на границы интервала $(-a, a)$ и

$$R(\xi)=\int\limits_0^{\nu_a} g[t,\ x_t,\ \Theta(t,\ x_t)]\,dt \tag{1}$$

— функционал, связанный с поведением процесса до первого выхода из интервала $(-a, a)$.

Наша задача состоит в том, чтобы в каждый момент времени t на основе наблюдаемых данных $\{\xi_\tau,\ \tau<t,\ |\xi_\tau|<a\}$ так изменять (регулировать) локальный снос $\theta\in\Theta$, чтобы среднее значение функционала (1) достигало своего минимума или максимума.

Хорошо известны затруднения (см., например, [1], [2]), связанные с общим определением правил регулирования при непрерывном наблюдении. Здесь мы будем, следуя [1], применять следующий подход. Пусть $\Delta=\frac{1}{2^n}$, $n\geqslant 0$, и регулирование может осуществляться в моменты времени, кратные Δ (так называемое согласно [3] кусочно-постоянное регулирование). Регулирующее правило δ^Δ зададим системой функционалов

$$\{\Phi^\Delta(t,\ \xi^t,\ \theta^t)\},\ t=0,\ \Delta,\ 2\Delta,\ \ldots;\ \xi^t=\{\xi_s,\ s\leqslant t\},\ \theta^t=\{\theta_s,\ s\leqslant t\},$$

которые при каждом t представляют собой распределение вероятностей для параметра θ в момент $t+\Delta$. Естественно, что эта система функционалов должна удовлетворять определенным условиям измеримости и согласованности (см., например, [4]). Тем самым, для правила δ^Δ определено математическое ожидание функционала (1)

$$r(\delta^\Delta)=\mathbf{M}_{\delta_\Delta} R(\xi).$$

Назовем правило регулирования δ регулярным, если существует такая последовательность $\{\delta^\Delta\}$, что $\lim\limits_{\Delta\to 0} r(\delta^\Delta)=r(\delta)$. Оптимальным правилом будем называть правило δ^*, для которого $r(\delta^*)$ является экстремумом $r(\delta)$ по всем регулярным правилам δ.

В том случае, когда $g[t, x_t, \theta_t]=1$, $\mathbf{M}R(\xi)$ — среднее время пребывания процесса в интервале $(-a, a)$ до первого выхода на границу. Если $\Theta = [-\bar\theta, \bar\theta]$, то из результата И. В. Гирсанова [5] вытекает, что оптимальное регулирование состоит в придании в момент t параметру θ „крайнего" значения $\bar\theta$, если $\xi_{t-0}<0$ и $-\bar\theta$, если $\xi_{t-0}>0$. Здесь важно отметить, что оптимальное регулирование с целью как можно дольше удержать процесс в интервале $(-a, a)$ оперирует только с крайними значениями параметра θ. Аналогичный результат верен и в более общей обстановке (см. [5]), если только предполагать, что функция $g(t, x_t, \theta_t)$ явно не зависит от параметра θ.

В настоящей заметке мы покажем, что в случае явной зависимости функции $g(t, x_t, \theta_t)$ от параметра θ оптимальное регулирование должно использовать, вообще говоря, весь спектр возможных значений параметра θ. По-видимому, в этом направлении можно доказать весьма общий результат, принципиальная же его возможность станет ясной из рассматриваемого нами частного случая функции $g(t, x, \theta)=1-c(\theta)$.

Если функцию $c(\theta)$ истолковывать как стоимость в единицу времени, когда регулируемый параметр равен θ, а 1 — как „доход" в единицу времени от пребывания ξ_t внутри $(-a, a)$, то $\mathbf{M}R(\xi)$ можно интерпретировать как среднее значение „чистого дохода" до первого достижения границ интервала $(-a, a)$. Относительно функции стоимости $c(\theta)$ будем предполагать, что она задана на множестве

$$\hat\Theta=\{\theta: 0,\ \underline{\theta}\leqslant|\theta|\leqslant\bar\theta,\ 0<\underline{\theta}<\bar\theta<\infty\},$$

неотрицательна, симметрична $(c(\theta)=c(-\theta)$, дважды дифференцируема для $\underline{\theta}<|\theta|<\bar\theta$ и $c(0)=0$. При интерпретации (1) как дохода естественно допустить, что $c'(\theta)\geqslant 0$, $\theta>0$. В этих предположениях мы покажем, что структура оптимальных правил определяется знаком второй производной $c''(\theta)$ функции стоимости. А именно, если $c''(\theta)>0$, т. е. стоимость растет довольно быстро, то надо использовать, вообще говоря, все значения $\theta\in\Theta$; если же $c''(\theta)\leqslant 0$, т. е. стоимость растет не слишком быстро, то оптимальное правило использует только значения $\theta=0$, $\pm\bar\theta$.

Этот результат в известном смысле обобщает результат [5]. Точные формулировки содержатся в приводимых ниже теоремах, в которых используются следующие обозначения:

$$\underline{\alpha}=\frac{c(\theta)}{2\underline{\theta}},\quad \bar\alpha=\frac{c(\bar\theta)}{\bar\theta}; \tag{2}$$

$\alpha(\theta)$ — решение дифференциального уравнения

$$\alpha'=\frac{c''(\theta)}{2[1-c(\theta)-\theta c'(\theta)]},\quad \alpha(\underline{\theta})=\alpha; \tag{3}$$

$B_i(\theta)$ — решения дифференциального уравнения

$$B'(\theta)=\frac{1-c(\theta)+\theta c'(\theta)}{\theta^3}\cdot[\theta\alpha(\theta)-1]$$

с начальными условиями

$$B_1(\bar\theta)=\frac{1}{2\bar\theta}\exp\left\{2\bar\theta\left[a-\int_{\underline{\theta}}^{\bar\theta}\alpha(\theta)\,d\theta\right]\right\}-\frac{1-c(\bar\theta)}{\bar\theta}\,a,$$

где $\alpha(\bar\theta)=a$,

и

$$B_2(\bar\theta)=\frac{1}{2\theta}\exp\left\{2\bar\theta\left[a-\int\limits_{\underline{\theta}}^{\bar\theta}\alpha(\theta)\,d\theta\right]\right\}-\frac{1-c(\theta)}{\bar\theta}\,a.$$

Теорема 1. *Если* $c''(\theta)>0$ *и* $\frac{c(\underline{\theta})}{\underline{\theta}}\leqslant c'(\underline{\theta})$, *то оптимальное правило* δ^* *состоит в непрерывном наблюдении координаты** x *процесса* ξ_t *и в следующем использовании в зависимости от* x *значений параметра* θ:

а) $\theta=0$ *при* $|x|<a$, *в том случае, когда* $\underline{\alpha}\geqslant a$;

б) $\theta=0$ *при* $|x|<\underline{\alpha}$, *и* $\theta=-(\operatorname{sign}x)\,\alpha^{-1}(|x|)$, *при* $\underline{\alpha}\leqslant x<a$, *в том случае, когда* $\underline{\alpha}\leqslant x<a$, $\alpha(\bar\theta)\geqslant a$;

в) $\theta=0$ *для* $|x|<\underline{\alpha}$, $\theta=-(\operatorname{sign}x)\,\alpha^{-1}(|x|)$ *для* $\underline{\alpha}\leqslant|x|<\alpha(\bar\theta)$, *и* $\theta=-(\operatorname{sign}x)\,\bar\theta$ *для* $\alpha(\bar\theta)\leqslant|x|<a$, *в том случае, когда* $\alpha(\bar\theta)<a$.

Соответствующие случаям а, Б, В, максимальные доходы равны:

$$\text{а) } r^*=a^2; \tag{4}$$

$$\text{б) } r^*=\left(\underline{\alpha}^2-\frac{1}{2\underline{\theta}^2}\right)+\frac{1-c(\underline{\theta})}{\underline{\theta}}+B_1(\underline{\theta}); \tag{5}$$

$$\text{в) } r^*=\left(\underline{\alpha}^2-\frac{1}{2\underline{\theta}^2}\right)+\frac{1-c(\underline{\theta})}{\underline{\theta}}+B_2(\underline{\theta}). \tag{6}$$

Теорема 2. *Если* $c''(\theta)\leqslant 0$ *и* $\frac{c(\underline{\theta})}{\underline{\theta}}\geqslant c'(\underline{\theta})$, *то оптимальное правило* δ^* *состоит в непрерывном наблюдении координаты* x *процесс* ξ_t *и в следующем использовании в зависимости от* x *значений параметра* θ

а) $\theta=0$, $|x|<a$, *если* $\bar\alpha\geqslant a$;

б) $\theta=0$, *при* $|x|<\bar\alpha$, *и* $\theta=-(\operatorname{sign}x)\,\bar\theta$, *при* $\bar\alpha\leqslant|x|<e$, *если* $\bar\alpha<a$.

Соответствующие максимальные доходы равны: в случае

а) $r^*=a^2$ *и в случае б)*

$$r^*=\bar\alpha^2+\frac{1}{2\bar\theta^2}\left\{\exp\left[2\bar\theta\,(a-\bar\alpha)\right]-1\right\}+\frac{c(\bar\theta)-1}{\bar\theta}\,(a-\bar\alpha). \tag{7}$$

Остановимся на основных моментах доказательства этих теорем. Прежде всего отметим, что вследствие симметрии функции стоимости $c(\theta)$ оптимальное правило симметричным образом зависит от координаты x процесса ξ_t, поэтому достаточно определить оптимальное правило для $x>0$, т. е. на интервале $(0,\ a)$. Предположим теперь, что множество Θ состоит из $2k+1$ элементов

$$\Theta^k=\{\theta:0,\ \pm\theta_1,\ \ldots,\ \pm\theta_k\},\quad 0<\theta_1<\ldots<\theta_k$$

и пусть

$$r(x)=\sup_{\{\delta\}}\mathrm{M}_\delta\left\{\int\limits_0^{\nu_a}[1-c(\theta_t)]\,dt/\xi(0)=x\right\},$$

$$r_i(x)=\sup\mathrm{M}_\delta\left\{\int\limits_0^{\nu_a}[1-c(\theta_t)]\,dt/\xi(0)=x,\theta(0)=-\theta_i\right\},\quad \theta_0=0,\ i=0,\ 1,\ \ldots,\ k.$$

* Легко видеть, что оптимальное правило в момент времени t не зависит явно от t.

Используя методы, развитые в работах [1], [2], нетрудно доказать следующую лемму.

Лемма. *Для того, чтобы* $r(x)$, $x>0$ *являлся максимальным доходом, необходимо, чтобы существовало разбиение интервала* $(0, a)$ *на* $n+1\leqslant k+1$ *интервалов** (α_i, α_{i+1})

$$0=\alpha_0<\alpha_1<\ldots<\alpha_n<\alpha_{n+1}=a, \tag{8}$$

внутри каждого из которых $r(x)$ *удовлетворяет уравнению*

$$\frac{1}{2}r''(x)-\theta_i r'(x)+1-c(\theta_i)=0,\ r(x)=r_i(x),\ i=0,\ 1,\ \ldots,\ n, \tag{9}$$

а на границах подчиняются соотношениям

$$r_i^{(l)}(\alpha_{i+1})=r_{i+1}^{(l)}(\alpha_{i+1})=r^{(l)}(\alpha_{i+1});\ l=0,\ 1,\ 2;\ i=0,\ 1,\ \ldots,\ n-1 \tag{10}$$

$$r(\alpha_{n+1})=0.$$

В [6] построены решения уравнений (9)

$$r_i(x)=A_i e^{2\theta_i x}+\frac{1-c(\theta_i)}{\theta_i}x+B_i,$$

где постоянные A_i, B_i и α_i находятся из условий (10). При этом оказывается, что

$$\alpha_1=\frac{c(\theta_1)}{2\theta_1},\quad \alpha_{i+1}=\alpha_i+\frac{1}{2\theta_i}\ln\frac{w_{i,i+1}}{w_{i-1,i}},\quad i=1,\ \ldots,\ n-1 \tag{11}$$

и

$$r(0)=\frac{c^2(\theta_1)}{4\theta_1^2}+\sum_{i=1}^{n-1}\frac{1}{2\theta_i^2}\left\{w_{i,i+1}-w_{i-1,i}-[1-c(\theta_i)]\ln\frac{w_{i,i+1}}{w_{i-1,i}}\right\}+$$
$$+\frac{w_{n-1,n}}{2\theta_n^2}\left\{\exp[2\theta_n(a-\alpha_n)]-1\right\}-\frac{1-c(\theta_n)}{\theta_n}(a-\alpha_n), \tag{12}$$

где

$$w_{i,i+1}=\frac{\theta_i c(\theta_{i+1})-\theta_{i+1}c(\theta_i)}{\theta_{i+1}-\theta_i}+1.$$

Вследствие (11) условие (8) можно записать в следующей форме

$$\frac{c(\theta_1)}{\theta_1}<\frac{c(\theta_2)-c(\theta_1)}{\theta_2-\theta_1}<\ldots>\frac{c(\theta_n)-c(\theta_{n-1})}{\theta_n-\theta_{n-1}}. \tag{8^1}$$

В [6] также показано, что доход может только увеличиться, если использовать еще одно такое значение параметра $\theta=\hat{\theta}$, что условие (8^1) выполнено и для $\theta_1,\ \ldots,\ \theta_n,\ \hat{\theta}$. Поскольку условия теоремы 1, как легко видеть, являются аналогом условий $(8')$ для $\Theta=\hat{\Theta}$, то надо использовать все $\theta\in\hat{\Theta}$. При этом формулы (5) и (6) получаются соответствующим предельным переходом из формул (11) и (12). В свою очередь условия теоремы 2 показывают, что условие $(8')$ верно только для одного значения параметра $\theta>0$, поэтому оптимальное правило должно состоять в использовании самое большее трех значений параметра: 0, $\pm\theta$, $\theta>0$, $\theta\in\hat{\Theta}$, причем легко доказать, что наибольший доход получается при использовании наибольшего значения $\theta=\bar{\theta}$. Формулы (2) и (7) – частный случай формул (11) и (12) при $n=1$ и

* Интервалов может быть меньше, чем $k+1$, так как некоторые значения параметра θ не используются; при этом, чтобы не писать громоздких индексов, нумерацию установим заново.

$\theta_1=\theta$. В качестве иллюстрации к теоремам 1 и 2 приведем следующих два примера:

Пример 1. Пусть $c(\theta)=\mu\theta^2$, $-\infty<\theta<\infty$. Согласно теореме 1 нетрудно показать, что оптимальное правило в этом случае состоит в придании параметру θ значений

$$\theta=-\frac{1}{\sqrt{\mu}}\operatorname{tg}\frac{x}{\sqrt{\mu}},\quad |x|<a,$$

где x — координата наблюдаемой точки. Соответствующий этому оптимальному правилу доход равен

$$r^*=\mu\ln\left(1+\operatorname{tg}^2\frac{a}{\sqrt{\mu}}\right).$$

Пример 2. Пусть $c(\theta)=\mu|\theta|$, $0\leqslant|\theta|\leqslant\bar{\theta}$, $\frac{\mu}{2}<a$. Оптимальное правило (см. теорему 2) состоит в использовании следующих значений параметра θ:

$$\theta=0 \text{ при } |x|<\frac{\mu}{2} \text{ и } \theta=-(\operatorname{sign}x)\,\theta \text{ при } \frac{\mu}{2}\leqslant|x|<a.$$

Соответствующий доход равен

$$r^*=\frac{\mu^2}{4}+\frac{1}{2\bar{\theta}}\left\{\exp\left[2\bar{\theta}\left(a-\frac{\mu}{2}\right)\right]-1\right\}-\frac{1-\mu\bar{\theta}}{\bar{\theta}}\left[a-\frac{\mu}{2}\right).$$

В заключение выражаю признательность А. Н. Ширяеву за ценные советы при решении данной задачи.

Поступила в редакцию
29.XI.1963

ЛИТЕРАТУРА

1. В. С. Михалевич. Последовательные байесовские решения и оптимальные методы контроля. Теория вероятн. и ее примен. 1, 4 (1956), 437—465.
2. Беллман Р. Динамическое программирование. ИЛ, 1960 г.
3. Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе, Е. Ф. Мищенко. Математическая теория оптимальных процессов. Физматгиз, 1961 г.
4. А. Н. Ширяев О задачах скорейшего обнаружения момента изменения вероятностных характеристик наблюдаемого процесса. Резюме доклада на заседании секции теории вероятностей и матем. статистики. Теория вероят. и ее примен. VII, 2 (1962), 236—238.
5. И. В. Гирсанов „Минимаксные задачи в теории диффузионных процессов", ДАН, т. 136. № 4 (1961 г.)
6. В. А. Колемаев. О статистическом регулировании размеров изделий серийного производства, Изв. АН СССР, Техническая кибернетика, № 4, 1963.

APIE VIENĄ VINERIO PROCESO OPTIMALAUS VALDYMO UŽDAVINĮ

V. A. KOLEMAJEVAS

(Reziumė)

Nagrinėjamas Vinerio proceso valdymo, reguliuojant lokalinį pernešimą $\theta\in\Theta$ tam, kad gautume funkcionalo (1) matematinės vilties ekstremumą, uždavinys.

Straipsnyje parodoma, kad, kai pointegralinė funkcija $g(t, x, \theta)$ priklauso tiesiogiai nuo parametro θ, optimali taisyklė, aplamai, turi panaudoti visas reikšmes $\theta\in\Theta$. Detaliai išnagrinėtas atvejis, kai pointegralinė funkcija lygi $g(t, x, \theta)=1-c(\theta)$.

ON A PROBLEM OF THE OPTIMAL CONTROL OF THE WIENER PROCESS

V. A. KOLEMAJEV (MOSCOW)

(Summary)

It is considered a problem of the optimal control of the Wiener process by means of variation of the parameter $\theta \in \Theta$ with an aim to give an extremal mathematical expectacion of the functional (1).

In the note it is shown that, when the function $g(t, x. \theta)$ depends on parameter θ evidently, an optimal rule, in generally, must employ all values $\theta \in \Theta$. In detail it is considered the case, when the function $g(t, x\ \theta) = 1 - c(\theta)$.