УДК 519.8; 368; 65.0

Б.В. НОРКИН

О СТОХАСТИЧЕСКОМ ОПТИМАЛЬНОМ УПРАВЛЕНИИ ПРОЦЕССАМИ РИСКА В ДИСКРЕТНОМ ВРЕМЕНИ

Изучается проблема стохастического оптимального управления дивидендной политикой страховой компании. Функционирование страховой компании описывается стохастическим процессом риска в дискретном времени с вычитанием дивидендов. Оптимизации подлежат средние суммарные дисконтированные дивиденды, по-лученные до момента разорения, плюс среднее дисконтированное время жизни процесса с некоторым коэффициентом. Установлены условия существования и единственности решения уравнений Беллмана для этой задачи.

Введение. Проблема стохастического оптимального управления дивидендной политикой страховой компании изучалась в работах [1 – 4]. В частности, в работе де Финетти [1] (см. также [2]) обнаружен следующий парадокс: при оптимальной стратегии управления (на бесконечном интервале времени) страховая компания разоряется с вероятностью единица. Данный результат показывает, что рассматриваемая постановка задачи не является полностью удовлетворительной. Желательно найти такую привлекательную дивидендную стратегию, при которой вероятность разорения мала.

В настоящей работе данная задача рассматривается как двухкритериальная: необходимо максимизировать суммарный доход (суммарные дисконтированные дивиденды) и минимизировать риск разорения. В качестве меры риска берется среднее (дисконтированное) время жизни компании. Дисконтирование означает, при принятии текущих решений год жизни в далеком будущем значит меньше, чем год жизни в ближайшем будущем. Дисконтирование также позволяет избежать формальной проблемы возможного бесконечного среднего времени жизни. В стохастическом программировании обычно используются другие меры риска, связанные с нарушением ограничений (в нашем случае – неотрицательность капитала компании) – это степень нарушения или вероятность нарушения (в нашем случае – ве-роятность разорения). В работе [5] формально показано, что при увеличении штрафа за нарушение ограничений вероятность нарушения стремится к нулю. В настоящей статье показатели доходности и риска агрегируются в один критерий: к средним дисконтированным дивидендам добавляется (с варьируемым коэффициентом) среднее дисконтированное время жизни. Для полученной агрегированной задачи оптимального управления оказывается справедлив принцип динамического программирования и выполнены уравнения Беллмана. Поскольку в данной задаче значения функции Беллмана разрывны и не ограничены, то стандартная методика доказательства существования и единственности решения уравнения Беллмана не применимы. Мы модифицируем эту методику, пользуясь оценками сверху и снизу для функции Беллмана. Решение задачи (функция Беллмана и оптимальное позиционное управление) находится методом последовательных приближений. Варьируя коэффициент агрегации, можно построить Парето-оптимальное множество значений показателей доходности и риска, а также соответствующие оптимальные управления. Далее решение выбирается субъективно из множества Парето-оптимальных точек. Когда управление (дивидендная политика) выбрана, то вероятность разорения может быть найдена из решения соответствующего интегрального уравнения [6] или оценена методом Монте-Карло [7]. Отметим, что задача оптимизации управления, заданного в параметрической форме как функция текущего капитала и конечномерных параметров, при ограничении на вероятность разорения может быть приближенно решена методом работы [8].

Задача стохастического оптимального управления дивидендной политикой страховой компании. Процесс риска описывает стохастическую эволюцию резервов страховой компании, предназначенных для покрытия страховых требований. Математическая модель эволюции резервов X^t в дискретном времени имеет вид [1, 2]:

$$X_{t+1} = f\left(X_t, U_t, Y_{t+1}\right) = X_t - U_t + Y_{t+1}, \quad X_0 = x, \quad U_t \in \left[0, X_t\right], \tag{1}$$

$\left\{Y_t\right\}$ – независимые одинаково распределенные (как Y) случайные величины с общей функцией распределения F. Обозначим момент остановки (разорения) $\tau = \sup\left\{t \in [0,T] : \min_{0\le t'<t} X_{t'} \ge 0\right\}$ и множество допустимых управлений $\bar{U} = \left\{U_t \in [0,x],\, t \in [0,T]\right\}$.

Среднее время жизни $\mathbf{E}\tau$ является не очень удобным индикатором риска, поскольку оно может равняться бесконечности. Поэтому наряду с $\mathbf{E}\tau$ мы будем рассматривать также так называемое среднее дисконтированное время жизни

$$T' = \mathbf{E}\sum\nolimits_{t=0}^{\tau-1}\gamma^t = \left(1-\mathbf{E}\gamma^{\tau}\right)\Big/(1-\gamma) \le 1/(1-\gamma).$$

Очевидно, $\left(1-\gamma^{\mathbf{E}\tau}\right)\Big/(1-\gamma) \le T' \le E\tau$ (если $E\tau < \infty$). Дисконтирование можно интерпретировать как наличие некоторого случайного фактора (с биноминальным распределением с параметром γ), который может остановить процесс риска независимо от его текущего состояния [9].

Функция выигрыша за один период времени $r(X_t, U_t)$ имеет вид:

$$r_\lambda(X_t, U_t) = \begin{cases} \lambda + U_t, & X_t \ge 0, \\ 0, & X_t < 0, \end{cases}$$

$\lambda \ge 0,$ а дисконтированная функция выигрыша за T периодов имеет вид:

$$V_T^U(x) = \mathbf{E}\left[\sum_{t=0}^{T} \gamma^t r(X_t, U_t)\right] = \mathbf{E}\sum_{t=0}^{\tau-1} \gamma^t (\lambda + U_t) = \lambda \mathbf{E}\sum_{t=0}^{\tau-1} \gamma^t + \mathbf{E}\sum_{t=0}^{\tau-1} \gamma^t U_t,$$

$0 < \gamma \le 1,$ $U = \{U_t, 0 \le t \le T\}.$ Определим функцию Беллмана

$$V_t(x) = \sup_{\{U: U_t \in [0,x]\}} V_t^U(x), \qquad 0 \le t \le T. \tag{2}$$

По определению, будем считать, что $V_t(x) = 0$ для $x < 0$. Однокритериальная задача (2) (т. е. с $\lambda = 0$) и дискретным распределением Y рассмотрена в [1, 2]. Двухкритериальная задача управления дивидендами в частном случае, когда Y принимает только два значения, анализировалась в [4].

Обоснование метода динамического программирования.

Лемма 1 (ограниченность функции Беллмана). Функция $V_t(x)$ удовлетворяет ограничениям

$$x + \lambda \le V_t(x) \le x + \left(\lambda + \gamma \mathbf{E}\max\{0, Y\}\right)/(1-\gamma).$$

Доказательство. Очевидно, $V_0(x) = x + \lambda$ и $V_t(x) \le V_{t+1}(x) \le V_T(x)$. В свою очередь,

$$V_T(x) \le \mathbf{E}\sum_{t=0}^{\tau-1} \gamma^t (\lambda + U_t) \le \lambda/(1-\gamma) + \mathbf{E}\sum_{t=0}^{\tau-1} \gamma^t U_t.$$

В силу [2, лемма 1.8], $\mathbf{E}\sum_{t=0}^{\tau-1} \gamma^t U_t \le \gamma \mathbf{E}\max\{0, Y\}/(1-\gamma)$. Лемма доказана.

Если функции Беллмана $V_t(\cdot)$ конечны, то при $x \ge 0$ они удовлетворяют рекуррентным соотношениям Беллмана [2, лемма 1.1]: $V_{-1}(x) \equiv 0,$

$$V_t(x) = \sup_{u \in [0,x]} \left\{ r_\lambda(x,u) + \gamma \mathbf{E} V_{t-1}\left(f(x,u,Y)\right) \right\} =$$

$$= \sup_{u \in [0,x]} \left\{ \lambda + u + \gamma \mathbf{E} V_{t-1}\left(x - u + Y\right) \right\}, \qquad 0 \le t \le T < \infty. \tag{3}$$

Лемма 2 (свойства функций Беллмана и существование оптимальных управлений при конечном временном горизонте $T < \infty$). Пусть $\mathbf{E}\max\{0, Y\} < \infty$. Тогда функции $\{V_t(x), -\infty < x < \infty, 0 \le t \le T < \infty\}$ не убывают по x (при фиксированном t) и монотонно возрастают по t (при фиксированном x), полунепрерывны сверху по x (и, следовательно, непрерывны справа). Кроме того функции $\mathbf{E} V_{t-1}(x - u + Y)$ полунепрерывны сверху по $(x \ge 0, u \ge 0),$ а функции

$$u_t^*(x) = \sup\left\{ v \in \arg\max_{u \in [0,x]} \left\{ \lambda + u + \gamma \mathbf{E} V_{t-1}(x - u + Y) \right\} \right\}$$

являются измеримыми по Борелю оптимальными управлениями для задач (2).

Доказательство. По определению, $V_{-1}(x) \equiv 0$,

$$V_0(x) = \sup_{0\le u\le x}\left\{r_\lambda(x,u)\right\} = \sup_{0\le u\le x}\left\{\lambda + u\right\} = \lambda + x, \quad u_0^*(x) = x.$$

По индукции, очевидно, следует, что функции $V_t(x)$ не убывают по x. Покажем, что последовательность функций $\{V_t(\cdot)\}$ монотонно возрастает (по t). Очевидно, для любого x выполнено $V_{-1}(x) \le V_0(x)$. Если $V_{t-1}(x) \le V_t(x)$ для любого x, то для любого x' и всех u имеет место

$$\lambda + u + \gamma \mathbf{E} V_{t-1}(x' - u + Y) \le \lambda + u + \gamma \mathbf{E} V_t(x' - u + Y).$$

Поэтому

$$V_t(x') = \sup_{u\in[0,x']}\left\{\lambda + u + \gamma \mathbf{E} V_{t-1}(x' - u + Y)\right\} \le$$
$$\le \sup_{u\in[0,x']}\left\{\lambda + u + \gamma \mathbf{E} V_t(x' - u + Y)\right\} = V_{t+1}(x').$$

Таким образом, при каждом фиксированном x' числовая последовательность $\left\{V_t(x'),\ t = -1, 0, 1, \ldots\right\}$ монотонно возрастает.

Покажем, что все $V_t(x)$ полунепрерывны сверху по x, а $u_t^*(x)$ являются оптимальными управлениями для задач (2). Очевидно, утверждение выполнено для $t = 0$. Действуя по индукции, предположим, что функция $V_{t-1}(x)$ полунепрерывна сверху. Тогда функция $V_{t-1}(x - u + y)$ полунепрерывна сверху по (x,u) при каждом y, а по лемме 1

$$V_{t-1}(x - u + y) \le x - u + y + \frac{\lambda + \gamma \mathbf{E}\max\{0,Y\}}{1-\gamma} \le x + \max\{0,y\} +$$
$$+ \frac{\lambda + \gamma \mathbf{E}\max\{0,Y\}}{1-\gamma},$$

$$\mathbf{E} V_{t-1}(x - u + y) \le x + \mathbf{E}\max\{0,y\} + \left(\lambda + \gamma \mathbf{E}\max\{0,Y\}\right)/(1-\gamma).$$

В силу леммы Фату [10, гл. II, §6, теорема 2(b)] функция $\mathbf{E} V_{t-1}(x - u + Y)$ полунепрерывна сверху по $(x \ge 0, u \ge 0)$. Отсюда следует, что функция максимума $V_t(x) = \sup_{u\in[0,x]}\left\{\lambda + u + \gamma \mathbf{E} V_{t-1}(x - u + Y)\right\}$ полунепрерывна сверху, а отображение $\bar{U}_t(x) = \arg\max_{u\in[0,x]}\left\{\lambda + u + \gamma \mathbf{E} V_{t-1}(x - u + Y)\right\}$ и функция $u_t^*(x) = \ = \max\left\{u \in \bar{U}_t(x)\right\}$ измеримы (по Борелю, см. [11, разд. 14]). Согласно [2, следствия 1.2] последовательность $U^*(x) = \left\{u_t(x), 0 \le t \le T\right\}$ является оптимальным управлением для задач (2),

$$V_t(x) = \sup_{\{U: U_t \in [0,x]\}} V_t^U(x) = V_t^{U^*}(x) = \mathbf{E}\sum_{t=0}^{\tau^*-1} \gamma^t(\lambda + u_t^*(x_t)), \qquad 0 \le t \le T,$$

где τ^* – момент остановки (разорения) процесса (1) с $U^* = \left\{U_t^* = u_t^*(x_t)\right\}$. Лемма доказана.

Случай бесконечного горизонта $T=\infty$. Уравнение Беллмана имеет вид

$$V(x)=\sup_{u\in[0,x]}\left\{r_\lambda(x,u)+\gamma\mathbf{E}V\left(f(x,u,Y)\right)\right\}. \tag{4}$$

Пусть $\{V_t(x)\}$ – монотонная последовательность (монотонных полунепрерывных сверху по лемме 2) функций (3).

Теорема 1. (Свойства функции Беллмана и существование оптимальных управлений при $T=\infty$). Пусть $\mathbf{E}\max\{0,Y\}<\infty$. Тогда предел $V(x)=\lim_{t\to\infty}V_t(x)$ существует, является монотонной полунепрерывной сверху функцией. Функция $V(x)$ – единственное полунепрерывное сверху решение уравнения (4), удовлетворяющее условию $x\le V(x)\le x+C_0$ для всех $x\ge 0$, где C_0 – произвольная константа. Функция $\varphi_x(u)=\mathbf{E}V(x-u+Y),\quad u\in[0,x]$, полунепрерывна сверху по (x,u), а функция

$$u^*(x)=\sup\left\{v\in\arg\max_{u\in[0,x]}\left\{\lambda+u+\gamma\mathbf{E}V(x-u+Y)\right\}\right\} \tag{5}$$

измерима по Борелю и является решением задачи (2) для $T=\infty$.

Доказательство. По лемме 2, при каждом фиксированном x последовательность $\{V_t(x),\ t=-1,0,1,...\}$ монотонно возрастает, и по лемме 1 она ограничена сверху и имеет предел $V(x)$. Так как по лемме 2 все функции $\{V_t(x)\}$ не убывают по x, то таким же свойством обладает и их предел $V(x)$.

Докажем, что $V(\cdot)$ полунепрерывна сверху. Заметим, что в силу леммы 1

$$0\le\sup_{x\ge0}\left[V_t(x)-V_{t-1}(x)\right]\le\sup_{x\ge0}\left[V_t(x)-x-\left(V_{t-1}(x)-x\right)\right]\le$$

$$\le 2\left(\lambda+\gamma\mathbf{E}\max\{0,Y\}\right)/(1-\gamma)=C.$$

В силу оптимальности управлений,

$$V_{t+1}(x)=\sup_{u\in\mathbf{U}(x)}\left\{\lambda+u+\gamma\mathbf{E}V_t\left(x-u+Y\right)\right\}=$$

$$=\lambda+u^*_{t+1}(x)+\gamma\mathbf{E}V_t\left(x-u^*_{t+1}(x)+Y\right),$$

$$V_t(x)=\sup_{u\in\mathbf{U}(x)}\left\{\lambda+u+\gamma\mathbf{E}V_{t-1}\left(x-u+Y\right)\right\}\ge$$

$$\ge\lambda+u^*_{t+1}(x)+\gamma\mathbf{E}V_{t-1}\left(x-u^*_{t+1}(x)+Y\right).$$

поэтому

$$0\le V_{t+1}(x)-V_t(x)\le\gamma\mathbf{E}\left[V_t\left(x-u^*_{t+1}(x)+Y\right)-V_{t-1}\left(x-u^*_{t+1}(x)+Y\right)\right]\le$$

$$\le\gamma\sup_{x'\ge0}\left[V_t(x')-V_{t-1}(x')\right],$$

$$\sup_{x\ge 0}\left[V_{t+1}(x)-V_t(x)\right]\le\gamma\sup_{x\ge 0}\left[V_t(x)-V_{t-1}(x)\right]\le\gamma^t\sup_{x\ge 0}\left[V_1(x)-V_0(x)\right]=C\gamma^t,$$
$$V(x)-V_t(x)=\sum\nolimits_{k=t}^{\infty}\left(V_{k+1}(x)-V_k(x)\right)\le C\sum\nolimits_{k=t}^{\infty}\gamma^k=C\gamma^t/(1-\gamma). \qquad (6)$$

Таким образом, последовательность $\{V_t(\cdot)\}$ равномерно сходится к $V(\cdot)$ со скоростью геометрической прогрессии. Так как $V_t(\cdot)$ полунепрерывны сверху, то и их равномерный предел $V(\cdot)$ является полунепрерывной сверху функцией.

Покажем, что предел $V(\cdot)$ удовлетворяет уравнению Беллмана (4). Для каждого фиксированного x рассмотрим функции:

$$v_{t-1,x}(u)=\lambda+u+\gamma\mathbf{E}V_{t-1}(x-u+Y), \qquad v_x(u)=\lambda+u+\gamma\mathbf{E}V(x-u+Y),$$

$u\in[0,x]$. В силу леммы 2 функции $v_{t-1,x}(u)$ полунепрерывны сверху по u. Кроме того, они равномерно сходятся к (полунепрерывной сверху) функции $v_x(u)$ на интервале $u\in[0,x]$, а именно, в силу (6) имеет место

$$\left|v_x(u)-v_{t-1,x}(u)\right|\le\gamma\mathbf{E}\left|V(x-u+Y)-V_{t-1}(x-u+Y)\right|\le C\gamma^t/(1-\gamma),$$
$$V(x)=\lim\nolimits_{t\to\infty}V_t(x)=\lim\nolimits_{t\to\infty}\sup\nolimits_{u\in[0,x]}v_{t-1,x}(u)=\sup\nolimits_{u\in[0,x]}\lim\nolimits_{t\to\infty}v_{t-1,x}(u)=$$
$$=\sup\nolimits_{u\in[0,x]}v_x(u)=\sup\nolimits_{u\in[0,x]}\left[\lambda+u+\gamma\mathbf{E}V(x-u+Y)\right].$$

Таким образом, функция $V(x)$ удовлетворяет уравнению Беллмана (4) и существует измеримая (по Борелю [11, разд. 14]) функция $u^*(x)$ (5).

Заметим, что для любого управления $U'=\{U'_t\in[0,x]\}$ и соответствующей траектории $\{X'_t=X'_{t-1}-U'_{t-1}+Y_t,\ X'_0=x,\ U'_{t-1}\in[0,X'_{t-1}],\ t\ge 0\}$ выполнено условие [2, (1.3)]:

$$\sup\nolimits_{U'}\mathbf{E}\sum\nolimits_{k=t}^{\infty}\left|r_\lambda(X'_k,U'_k)\right|\gamma^k\le\frac{\lambda+\gamma\mathbf{E}\max\{0,Y\}}{1-\gamma}\sum\nolimits_{k=t}^{\infty}\gamma^k=\frac{\lambda+\gamma\mathbf{E}\max\{0,Y\}}{(1-\gamma)^2}\gamma^t.$$

Поэтому в силу [2, следствия 1.3] $u^*(x)$ является оптимальным управлением таким, что $V(x)=\sup_{\{U:U_t\in[0,x]\}}V^U(x)=V^{U^*}(x)=\mathbf{E}\sum\nolimits_{t=0}^{\tau^*-1}\gamma^t(\lambda+u^*(x_t))$, где $\left\{x^*_t=x^*_{t-1}-u^*(x^*_{t-1})+Y_t,\ x^*_0=x,\ t\ge 0\right\}$, τ^* – момент остановки процесса.

Докажем единственность решения $V(x)$. Пусть $V'(x)$ – другое полунепрерывное сверху решение уравнения (4), удовлетворяющее условию $x\le V(x)\le\ \le x+C_0$. Тогда функция $\varphi'_x(u)=\mathbf{E}V'(x-u+Y)$ полунепрерывна сверху по $u\in[0,x]$ и существует борелевская функция

$$u'(x)=\max\left\{v\in\arg\max\nolimits_{u\in[0,x]}\left\{\lambda+u+\gamma\mathbf{E}V'(x-u+Y)\right\}\right\}.$$

Справедливы соотношения:

$$V(x) = \sup_{u\in \mathbf{U}(x)} \left\{ \lambda + u + \gamma \mathbf{E} V\left(x-u+Y \right) \right\} = \lambda + u^*(x) + \gamma \mathbf{E} V\left(x - u^*(x) + Y \right),$$

$$V'(x) = \sup_{u\in \mathbf{U}(x)} \left\{ \lambda + u + \gamma \mathbf{E} V'\left(x-u+Y \right) \right\} = \lambda + u'(x) + \gamma \mathbf{E} V'\left(x - u'(x) + Y \right),$$

$$V(x) - V'(x) \le \gamma \mathbf{E}\Big(V(x-u^*(x)+Y) - V'(x-u^*(x)+Y) \Big) \le \gamma \sup_{x\ge 0} \left| V(x) - V'(x) \right|,$$

$$V'(x) - V(x) \le \gamma \mathbf{E}\big(V'(x-u'(x)+Y) - V(x-u'(x)+Y) \big) \le \gamma \sup_{x\ge 0} \left| V'(x) - V(x) \right|,$$

$$\sup_{x\ge 0} \left| V(x) - V'(x) \right| \le \gamma \sup_{x\ge 0} \left| V(x) - V'(x) \right|.$$

Следовательно, $\sup_{x\ge 0} \left| V(x) - V'(x) \right| = 0.$ Теорема доказана.

Заключение. В работе исследована двухкритериальная задача стохастического оптимального управления дивидендной политикой страховой компанией с учетом риска разорения в дискретном времени при общем распределении страховых требований. В задаче максимизируется свертка двух критериев, средних суммарных дисконтированных дивидендов и среднего (дисконтированного) времени жизни компании. Для свертки обоснован принцип динамического программирования, доказано существование и единственность решения уравнений Беллмана, а также существование оптимальных позиционных управлений.

Б.В. Норкін

ПРО СТОХАСТИЧНЕ ОПТИМАЛЬНЕ КЕРУВАННЯ ПРОЦЕСАМИ РИЗИКУ В ДИСКРЕТНОМУ ЧАСІ

Вивчається проблема стохастичного оптимального керування дивідендною політикою страхової компанії. Функціонування страхової компанії описується стохастичним процесом ризику в дискретному часі з відніманням дивідендів, а оптимізації підлягають середні сумарні дисконтовані дивіденди, отримані до моменту розорення, плюс середній дисконтований час життя процесу з деяким коефіцієнтом. Встановлено умови існування та єдиності розв'язку рівняння Беллмана для цієї задачі.

B.V. Norkin

ON STOCHASTIC OPTIMAL CONTROL OF THE RISK PROCESS IN DISCRETE TIME

The problem of stochastic optimal control of the dividend policy of an insurance company is studied. Functioning of the insurance company is described by a stochastic process risk in discrete time with deduction of dividends. Total (up to ruin) expected discounted dividends plus expected lifetime (with coefficient) are maximized. Conditions for existence and uniqueness of the solution of the Bellman equation for this problem are establish.

1. *De Finetti B.* Su un' impostazione alternativa della teoria collettiva del rischio // Transactions of the XV-th International Congress of Actuaries 2. – 1957. – P. 433 – 443.
2. *Schmidli H.* Stochastic control in insurance. – London: Springer-Verlag, 2008. – 254 p.
3. *Gerber H.U.* An introduction to mathematical risk theory. – Philadelphia: Huebner Foundation Monographs, 1979. – 164 c.
4. *Пиуновский А.Б.* Оптимальное управление случайными последовательностями в задачах с ограничениями. – М.: Научная книга, 1996. – 294 с.
5. *Ermoliev Y.M. et al.* Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks // Annals of Operations Research. – 2000. – Vol. 99. – P. 207 – 225.
6. *Норкин Б.В*. О решении основного интегрального уравнения актуарной математики методом последовательных приближений // Український математичний журнал. – 2007. – Том 59, № 12. – С. 112 – 127.
7. *Норкин Б.В*. Системный имитационный анализ и оптимизация страхового бизнеса // Кибернетика и системный анализ. – 2014. – № 2. – С. 112 – 125.
8. *Норкин Б.В*. Математические модели оптимизации страхового дела // Кибернетика и системный анализ. – 2011. – № 1. – С. 128 – 145.
9. *Ermoliev Y. et al.* Discounting, catastrophic risks management and vulnerability modeling // Mathematics and Computers in Simulation. – 2008. – Vol. 79. – P. 917 – 924.
10. *Ширяев А.Н.* Вероятность. – М.: Наука, 1980. – 574 с.
11. *Rockafellar R.T., Wets R.J-B.* Variational analysis. – Berlin, Heidelberg: Springer, 2004. – 734 p.

Получено 11.02.2014