

ИСПОЛЬЗОВАНИЕ ДВУХПороГОВОЙ СТРАТЕГИИ УПРАВЛЕНИЯ В СЛУЧАЙНОЙ СРЕДЕ С НОРМАЛЬНО РАСПРЕДЕЛЕННЫМИ ДОХОДАМИ

Лазутченко А.Н.¹

¹ФГБОУ ВПО «Новгородский государственный университет имени Ярослава Мудрого», Великий Новгород, Россия (173003, г. Великий Новгород, ул. Большая Санкт-Петербургская, 41), e-mail: aleskey@hotmail.ru

Рассмотрена пороговая стратегия управления в случайной среде с двумя действиями с нормально распределенными доходами. Цель управления задана в минимаксной постановке и состоит в минимизации математического ожидания максимальных потерь полного дохода и понижении суммарных потерь на некотором допустимом множестве параметров среды. С помощью написанной компьютерной программы среда исследуется на суммарные потери полного дохода. После этого суммарные потери вычисляются для двух порогов, оптимальные значения которых определяются после полного исследования всех их допустимых значений. Таким образом, в статье показано, что с помощью двухпороговой стратегии управления в случайной среде можно значительно снизить суммарные потери полного дохода на некотором множестве параметров среды.

Ключевые слова: случайная среда, двухпороговая стратегия управления, нормально распределенные доходы, минимакс.

USING THE TWO THRESHOLD MANAGEMENT STRATEGY IN A STOCHASTIC ENVIRONMENT WITH NORMAL DISTRIBUTED INCOMES

Lazutchenko A.N.¹

¹Yaroslav-the-Wise Novgorod State University, Velikiy Novgorod, Russia (173003, Velikiy Novgorod, street B. Sankt-Peterburgskaya, 41), e-mail: aleskey@hotmail.ru

Considered the threshold control strategy in a stochastic environment with two actions with normally distributed incomes. The purpose of the control is set to minimax formulation and is to minimize the mean of the maximum losses of total income and lowering the total loss on an allowable set of environmental parameters. With the help of a developed computer program environment investigated a total loss of total income. Thereafter, the total losses are calculated for the two thresholds, the optimal values are determined after examining all their allowable values. Thus, we have shown that using the two threshold management strategy in a random environment can significantly reduce the total losses of total income on a set of environmental parameters.

Keywords: random environment, the two threshold management strategy, normal distributed income, Minimax.

Введение

По аналогии с определением бинарной случайной среды, приведенным в [4], случайная среда (однородный процесс с независимыми приращениями в терминологии [5]) с нормально распределенными доходами – это управляемый случайный процесс ξ_t , значения которого интерпретируются как доходы, зависящие только от выбираемых в текущие моменты времени действий и имеющие нормальные распределения с плотностями

$$f(x | m_l) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-m_l)^2}{2}}, \quad (1)$$

если $y_t = l (l = 1, 2)$. При такой постановке задачи случайная среда описывается вектором математических ожиданий $\theta = (m_1, m_2)$. В данном случае параметр фиксирован, но неизвестен лицу, осуществляющему управление.

Постановка задачи

Введем целевую функцию потерь $L_T(\theta, \sigma)$, значениями которой являются потери за время моделирования, где σ – используемая стратегия. Если параметр θ известен, то наилучшей стратегией является та, которая применяет только то действие, которому соответствует большая из величин m_1, m_2 , и максимальный полный доход в этом случае равен $\max(m_1, m_2) \cdot T$. Если же θ неизвестен, то неизбежно возникают потери вследствие неполноты информации о среде, равные:

$$L_T(\theta, \sigma) = \max(m_1, m_2) \cdot T - E_{\sigma, \theta} \left(\sum_{t=1}^T \xi_t \right). \quad (2)$$

Здесь $E_{\sigma, \theta}$ представляет собой математическое ожидание потерь полного дохода. Предполагается, что ограничения на множество допустимых значений параметра θ имеют следующий вид:

$$\begin{aligned} \Theta &= \{(m_1, m_2) : |m_1 - m_2| \leq 2c_1\}, \\ 0 &< c_1 < \infty, 0 < c_2 < \infty. \end{aligned} \quad (3)$$

При использовании минимаксного подхода, предложенного, например, в [2], цель управления состоит в минимизации величины потерь полного дохода на множестве параметров Θ по множеству стратегий Σ . При этом минимаксный риск $R_T^M(\Theta)$ выглядит следующим образом:

$$R_T^M(\Theta) = \inf_{\Sigma} \sup_{\Theta} L_T(\sigma, \theta). \quad (4)$$

Для реализации этой цели предлагается использовать пороговую стратегию, предложенную в [6].

Стратегия управления с одним порогом

Итак, рассмотрим пороговую стратегию σ . Она применяет действия y_1 и y_2 среды по очереди, накапливая доходы X_1 и X_2 соответственно. На каждом шаге вычисляется абсолютная разность доходов на действиях $|X_1 - X_2|$. Действия применяются до тех пор, пока эта величина не превысит порога $\alpha \cdot T^{\frac{1}{2}}$, где T – полное время управления, α – пороговая константа, или не истечет время управления. Если время управления не истекло, то действие, которому соответствует меньшая величина набранного дохода, исключается из рассмотрения, а оставшееся время применяется только другое действие.

Можно показать, что наибольшие потери полного дохода будут иметь место при

$$m_1 = 0, m_2 = -\frac{\beta}{T^2}. \quad (5)$$

На основе пороговой стратегии σ , рассмотренной выше, была разработана программа. Прежде сделаем замечание. Целевая функция потерь $L_T(\theta, \sigma)$, вообще говоря, зависит от параметров σ и θ . Но для расчетов нам удобнее полагать, что она зависит от α и β , где α – пороговая константа, используемая пороговой стратегией, β – параметр среды.

Итак, алгоритм работы программы построен таким образом, что в ней для каждой пары $(\alpha; \beta)$ вычисляется средний доход $MZ_T = \sum_{i=1}^N Z_{i,T} / N$, где $Z_{i,T}$ – доход за одно моделирование, N – количество моделирований. Затем вычисляются средние потери математического ожидания дохода $L_T(\alpha, \beta) = m_1 \cdot T - MZ_T$. После этого при каждой константе α определяются минимальные потери $M_T(\beta) = \min_{\alpha} L_T(\alpha, \beta)$. При каждой константе β подбираются максимальные потери $R_T(\alpha) = \max_{\beta} L_T(\alpha, \beta)$. Точка, в которой $M_T(\beta) = R_T(\alpha)$, является минимаксной точкой, в которой достигается минимальная гарантированная величина потерь полного дохода.

В результате вычислений оптимальное значение α оказалось равным 0,55, при этом максимальные потери $L_T(\alpha, \beta) = 0,761$ соответствуют $\beta = 4$. Время моделирования выбиралось из условия $T = 100$, количество моделирований $N = 1\,000\,000$, что позволяет говорить о точности вычислений $\delta = 0,001$ [1]. Анализируя полученные результаты, видно, что параметры случайной среды с нормально распределенными доходами идентичны параметрам бинарной случайной среды [4], а потери примерно в 2 раза выше. Это соответствует теории.

Таблица 1 показывает результаты итоговых вычислений. Все потери в таблице являются приведенными путем деления на $T^{\frac{1}{2}}$. Желтым цветом обозначены локальные минимумы по α для каждого β , зеленым – локальные максимумы по β для каждого α , сиреневым – точка, в которой минимум по α равен максимуму по β .

Таблица 1 – Значения $L_T(\alpha, \beta)$, $\alpha \in (0; 1)$, $\beta \in (1; 10)$

$\beta \backslash \alpha$	0,43	0,46	0,49	0,52	0,55	0,58	0,61	0,64	0,67	0,7
2,0	0,652	0,645	0,639	0,637	0,634	0,637	0,638	0,641	0,644	0,651
2,5	0,721	0,713	0,706	0,702	0,698	0,700	0,704	0,709	0,713	0,720
3,0	0,766	0,756	0,744	0,740	0,738	0,739	0,743	0,749	0,756	0,767

3,5	0,788	0,774	0,764	0,758	0,756	0,761	0,764	0,772	0,783	0,793
4,0	0,793	0,779	0,766	0,762	0,761	0,765	0,771	0,782	0,794	0,806
4,5	0,787	0,772	0,762	0,758	0,758	0,760	0,770	0,781	0,796	0,812
5,0	0,775	0,758	0,749	0,745	0,748	0,755	0,764	0,777	0,795	0,811
5,5	0,756	0,740	0,735	0,731	0,735	0,741	0,754	0,769	0,788	0,807
6,0	0,733	0,721	0,714	0,715	0,722	0,731	0,745	0,761	0,781	0,803
6,5	0,710	0,702	0,697	0,700	0,708	0,717	0,736	0,755	0,775	0,797

Двухпороговая стратегия управления

Рассмотрим теперь потери на множестве допустимых параметров β для следующих двух случаев. На диаграмме линия 1 показывает минимальные потери для минимальных α для каждого β , линия 2 – потери для фиксированного $\alpha = 0,55$ (все потери также являются приведенными).

Как нетрудно заметить, начиная с $\beta = 0$ и до $\beta \approx 7$ потери за моделирование практически идентичны. После этого потери для случая с фиксированным α увеличиваются и в определенный момент устанавливаются на величине порога $\alpha \cdot T^{1/2}$. Очевидно, это связано с тем, что при больших значениях параметра β лучшее действие определяется достаточно быстро (так как, например, при $\beta = 10$ математические ожидания на действиях различаются значительно, $m_1 = 0, m_2 = -1$), но большое значение пороговой константы α (при $\alpha = 0,55$ порог $\alpha \cdot T^{1/2} = 5,5$) требует значительного времени на достижение этого порога. Этот недостаток стратегии можно устранить, если ввести дополнительный порог $\alpha_{th} < \alpha$, с помощью которого при больших β оптимальное действие будет определяться быстрее.

В

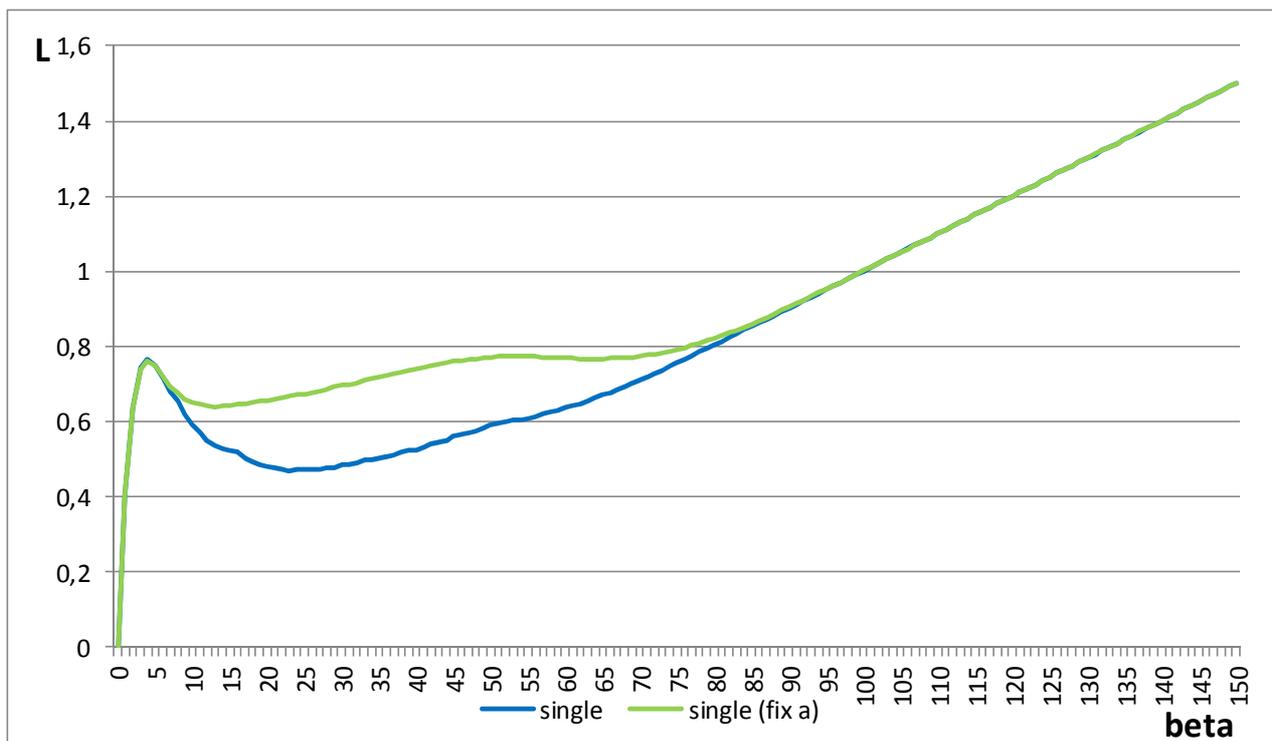


Рисунок 1 – Графики значений функции потерь на множестве значений β . Линия 1 показывает потери при минимальных α для каждого β , линия 2 – потери для фиксированного $\alpha = 0,55$.

этом случае стратегия будет выглядеть следующим образом. Начиная с некоторого времени $t = T_{th}$ текущий порог α_{th} заменяется на α , после чего моделирование продолжается в обычном порядке.

Очевидно, что теперь при высоких значениях β неоптимальное действие будет в среднем исключаться из рассмотрения быстрее. Введем новую величину $S = \int_{\beta} L(\alpha, \beta) d\beta$, которая будет показывать суммарные потери на всём множестве β . Рассматривание данной величины вполне уместно, так как можно считать, что значения множества β распределены равномерно.

Расчет значений

Итак, рассчитаем значения S для приведенных выше данных. Более конкретно, нас интересуют следующие значения: $S_{minimax}$ – для первого случая, S_{fix} – для случая фиксированного α и S_{th} – для случая двойного порога. Вычисления показывают, что $S_{minimax} = 47,24$, $S_{fix} = 57,67$. Далее был проведен полный расчет значений суммарных потерь за время моделирования при $T = 100$, $N = 1\,000\,000$, $\alpha = 0,55$, $\beta = 1, \dots, 150$, $a_{th} = 0,1 \dots 1$, $T_{th} = 10, \dots, 100$. Искались такие параметры α_{th} и T_{th} , при которых S_{th} будет наименьшим, при этом ни одно из значений потерь не превышает вычисленной величины гарантированных потерь полного дохода $L_T = 0,761$. В итоге удалось достигнуть значения $S_{th} = 47,35$ при

$T_{th} = 70$, $a_{th} = 0,35$. Диаграмма это демонстрирует, потери при таких параметрах обозначены на диаграмме:

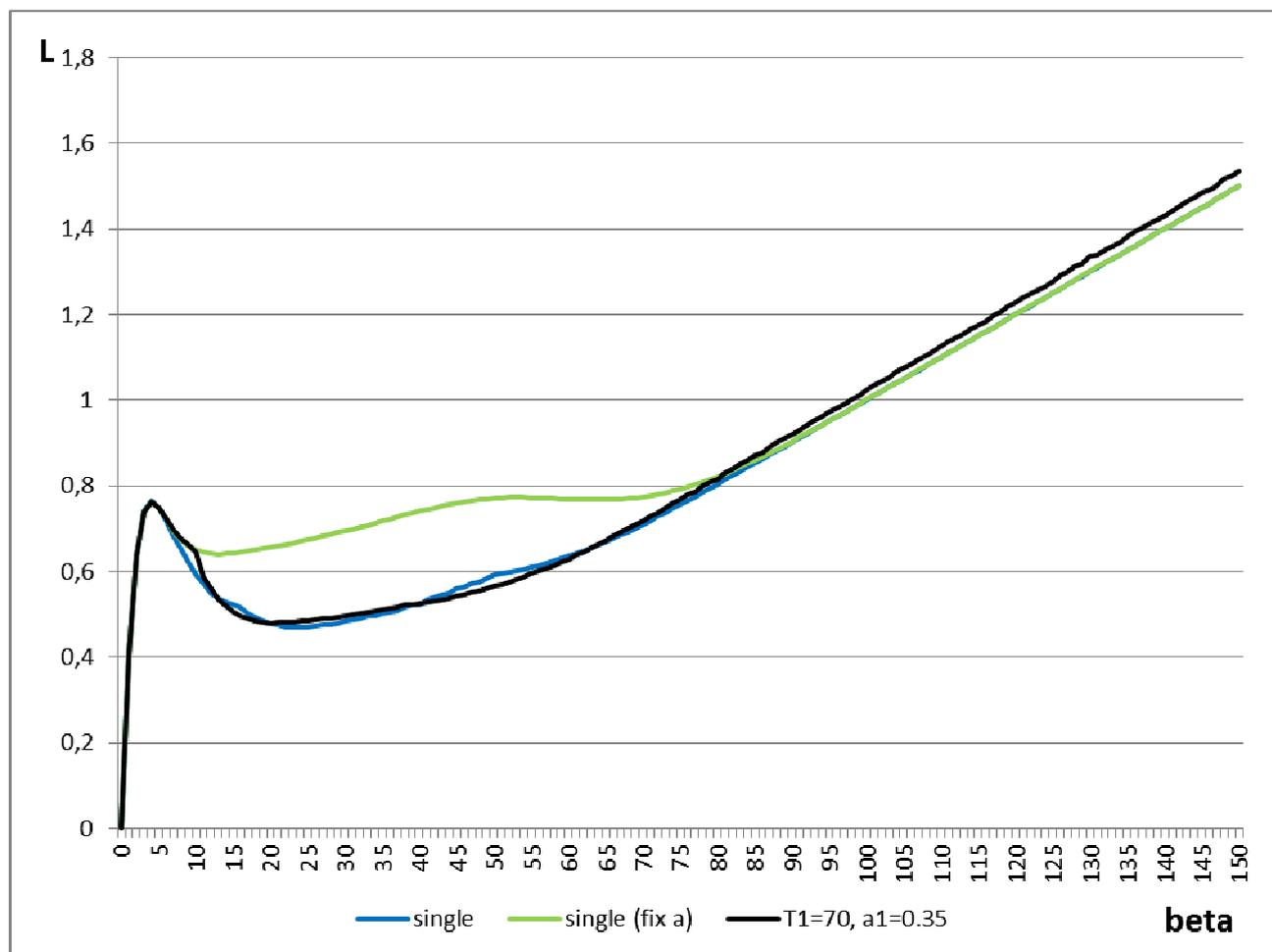


Рисунок 2 – Графики значений функции потерь на множестве значений β . Линия 1 показывает потери при минимальных α для каждого β , линия 2 – потери для фиксированного $\alpha = 0,55$, линия 3 – потери при фиксированном $\alpha = 0,55$, $\alpha_{th} = 0,25$, $T_{th} = 2000$.

Это позволяет судить о том, что введение дополнительного порога позволяет уменьшить потери на 22% по сравнению с фиксированным α , делая их сравнимыми с потерями первого случая.

Заклучение

Итак, рассмотрена пороговая стратегия управления в случайной среде с нормально распределенными доходами с двумя действиями. Найдены оптимальные значения пороговой константы и параметра среды. Также рассмотрены суммарные потери дохода на множестве допустимых параметров среды и показано, что их можно значительно снизить с помощью введения дополнительного порога.

Автор благодарит профессора А.В. Колногорова за помощь в постановке задачи и обсуждение полученных результатов.

Список литературы

1. Метод статистических испытаний (Метод Монте-Карло) / Н.П. Бусленко [и др.] – М. : Физматгиз, 1962.
2. Колногоров А.В. Нахождение минимаксных стратегий и риска в случайной среде (задача о двуруком бандите) // Автоматика и телемеханика [В. Новгород]. - 2011. - № 5.
3. Колногоров А.В., Шелонина Т.Н. Об инвариантности функции потерь для пороговой стратегии поведения в случайной среде // Вестн. Новг. гос. ун-та. - 2006. - № 39 - С. 18-21.
4. Лазутченко А.Н. Использование двухпороговой стратегии управления в бинарной случайной среде // Современные проблемы науки и образования. – 2013. – № 3. - URL: www.science-education.ru/109-9552 (дата обращения: 21.03.2014).
5. Срагович В.Г. Адаптивное управление. - М. : Наука, 1981. – 384 с.
6. Vogel W. // Ann. Math. Statist. - 1960. - V. 31. - P. 444–451.

Рецензенты:

Кириянов Б.Ф., д.т.н., профессор, ФГБОУ ВПО «Новгородский государственный университет имени Ярослава Мудрого», г. Великий Новгород.

Едемский В.А., д.ф.-м.н., профессор кафедры ПМИ, ФГБОУ ВПО «Новгородский государственный университет имени Ярослава Мудрого», г. Великий Новгород.