

ПОЛУЧЕНИЕ УПРОЩЕННЫХ СТРАТЕГИЙ ДЛЯ ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ
С РАЗЛИЧНЫМИ РАЗМЕРАМИ ГРУПП В СЛУЧАЙНОЙ СРЕДЕ

А.О.Олейников

PRODUCING ROUGH STRATEGIES FOR PARALLEL PROCESSING WITH DIFFERENT
GROUPS' SIZES IN RANDOM ENVIRONMENT

A.O.Oleinikov

Институт электронных и информационных систем НовГУ, duke.gh@yandex.ru

Предложены методы поиска упрощенных стратегий для задачи параллельного управления в случайной среде с различными размерами групп. Упрощенные стратегии находятся на основе точных при помощи минимизации квадратов отклонений от точной стратегии, либо минимизации ожидаемых потерь при применении упрощенной стратегии. Приведены результаты численных экспериментов.

Ключевые слова: управление в случайной среде, задача о двуруком бандите, параллельное управление

Parallel control in random environment with different groups' size was improved. Some methods for rough strategy finding were proposed. The first method is based on minimizing sum of deviation squares of the rough and accurate strategy. The second method is based on minimizing expected losses by numerical optimization. The results of the method application are given.

Keywords: control in random environment, two-armed-bandit problem, parallel control

1. Введение

Рассматривается задача об управлении в случайной среде (аналогичная задачам, описанным в [1-3]) в следующей постановке: пусть ξ_n , $n=1, \dots, N$ — управляемый случайный процесс, значения которого интерпретируются как доходы. При этом значения ξ_n зависят только от выбираемых в текущие моменты времени действий $\eta_n = \ell$ ($\ell=1,2$) и имеют нормальные распределения с математическими ожиданиями m_ℓ и единичными дисперсиями. Такая среда описывается векторным параметром $\theta=(m_1, m_2)$. Цель управления неформально можно описать как «максимизация суммарного ожидаемого дохода».

Если параметр θ известен, то следует на каждом шаге применять действие, которому соответствует большая из величин m_1 , m_2 ; тогда полный ожидаемый доход равен $N(m_1 \vee m_2)$. Функция

$$L_N(\sigma, \theta) = E_{\sigma, \theta} \left(\sum_{n=1}^N ((m_1 \vee m_2) - \xi_n) \right)$$

характеризует потери дохода вследствие неполноты информации, если параметр θ неизвестен. Здесь $E_{\sigma, \theta}$ обозначает математическое ожидание по мере, порожденной стратегией управления σ и параметром θ . Множество допустимых значений параметра имеет вид: $\Theta = \{(m_1, m_2) : m_1 - m_2 \leq 2c\}$, где c — некоторая константа ($0 < c < \infty$).

Будем использовать минимаксный подход. Величина $R_N^M(\Theta) = \inf_{\Sigma} \sup_{\Theta} L_N(\sigma, \theta)$ называется минимаксным риском, а соответствующая стратегия (если она

существует) — минимаксной стратегией (здесь Σ — множество всех стратегий).

В работе [4] показано, что минимаксные стратегия и риск могут быть найдены как байесовские, соответствующие наихудшему априорному распределению, а также даны рекуррентные формулы для их нахождения. Стратегию, найденную при помощи этих формул, будем далее называть точной.

2. Стратегии

Стратегия, найденная при помощи описанных в [4] рекуррентных формул, имеет пороговый характер и может быть представлена при помощи треугольной матрицы $A=(a_{ij})$. Значения i и j будут обозначать количество этапов управления, на которых были выбраны первое и второе действие соответственно. Тогда a_{n, n_2} — это пороговое значение для $Z = X_1 n_2 - X_2 n_1$, где X_ℓ — суммарный доход, полученный за все шаги, на которых было выбрано действие ℓ , а n_ℓ — количество таких шагов. Если пороговое значение меньше Z для текущего состояния, стратегия предписывает выбрать первое действие, в противном случае — второе.

Для визуализации данная матрица может быть представлена в виде набора графиков (по одному графику для каждого значения $n_1 + n_2$). Следуя [4],

введем обозначения $t = \frac{n_1 + n_2}{N}$, $t_\ell = \frac{n_\ell}{N}$. Выберем для оси абсцисс значения t_1/t , а для оси ординат — граничное значение $T(t_1, t_2) = Z \cdot N^{-3/2}$. График будет задан набором точек, определяемых пороговыми значениями Z для соответствующего шага. На рис.1 показан вид такого набора графиков для $N=15$ (показана только часть графиков).

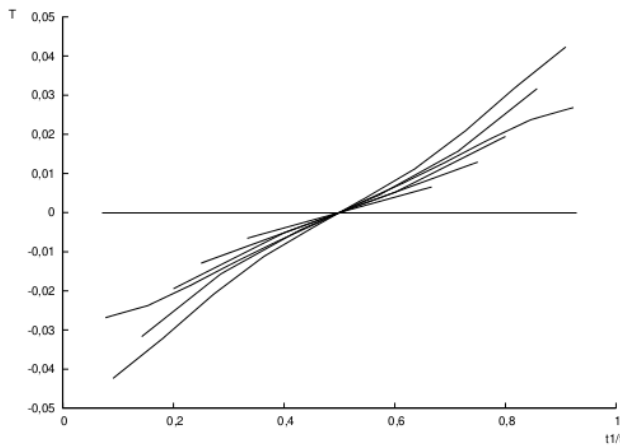


Рис.1. Графическое представление стратегии

Несложно проверить, что выполняется следующая лемма:

Лемма 1. Выполняются следующие равенства: $R_{n_1, n_2}^{(1)}(Z) = R_{n_2, n_1}^{(2)}(-Z)$; $R_{n_1, n_2}^{(2)}(Z) = R_{n_2, n_1}^{(1)}(-Z)$, где $R_{n_1, n_2}^{(\ell)}(Z)$ — потери в случае если при данных значениях n_1, n_2, Z будет выбран сначала ℓ -й вариант, а затем управление будет вестись оптимально.

3. Упрощенные стратегии

Из леммы 1 следует, что каждый график симметричен относительно точки $(0.5; 0)$. Используя это свойство, будем искать стратегии, которые могут быть табулированы не более чем двумя точками для каждого шага. Таким образом, каждый график упрощенной стратегии будет содержать не более пяти точек: две табулированные, две симметричные им и $(0.5; 0)$. Обозначим через x_i, y_i координаты точек точной стратегии, а через x_i^r, y_i^r — упрощенной.

Предлагаются следующие два способа получения упрощенных стратегий: в первом будем минимизировать сумму квадратов отклонений пороговых значений упрощенной стратегии от пороговых значений точной, а во втором будем минимизировать значение максимальных ожидаемых потерь, полученных при применении упрощенной стратегии.

Для обоих способов можно создавать три различных упрощенных стратегии, различающиеся количеством оптимизируемых параметров:

1) кодируем стратегию для каждого шага при помощи одной точки, при этом нам важен только угол наклона α графика такой стратегии;

2) кодируем стратегию для каждого шага при помощи двух точек, при этом значение x_1^r будем брать фиксированное. В таком случае мы оптимизируем два параметра: угол наклона α графика до точки x_1^r и угол β после нее;

3) кодируем стратегию для каждого шага при помощи двух точек, и при этом оптимизируем обе координаты x_1^r и y_1^r . Таким образом, у нас есть по три параметра для оптимизации на каждом шаге — x_1^r, α и β .

Таким образом, мы получаем 6 различных упрощенных стратегий на основе одной точной. Пример упрощенной стратегии, заданной одной точкой, показан на рис.2 пунктирной линией. На рис.3 показан пример упрощенной стратегии, заданной двумя точками.

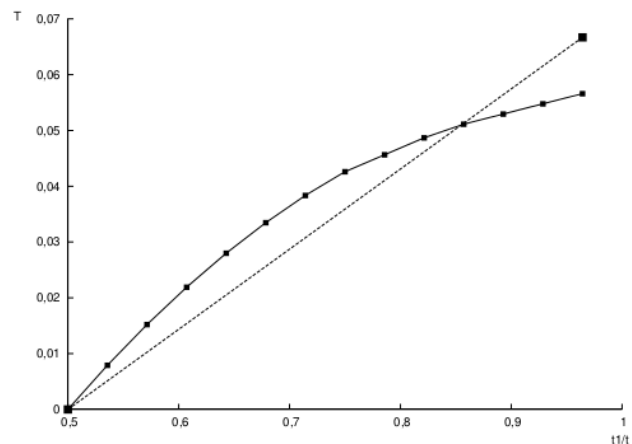


Рис.2. Часть точной стратегии и соответствующая ей упрощенная (заданная одной точкой)

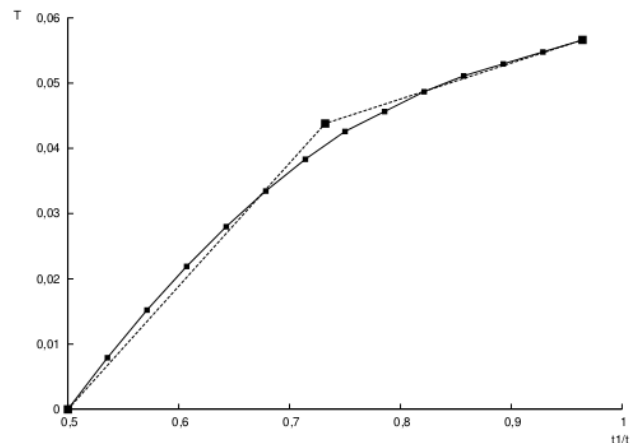


Рис.3. Часть точной стратегии и соответствующая ей упрощенная (заданная двумя точками)

Рассмотрим подробнее метод, основанный на минимизации суммы квадратов отклонений. В данном случае мы получаем упрощенные стратегии для каждого шага отдельно. Если мы оптимизируем только один параметр α , то нам достаточно обеспечить минимизацию функции $f(\alpha) = \sum_{i: x_i \in [0; 1]} (y_i - \alpha(x_i - 0.5))^2$.

Минимум такой функции будет достигаться в точке, в которой производная по α будет равняться нулю. Таким образом, получаем следующее выражение для вычисления α :

$$\alpha = (x_1^r - 0.5) \cdot \frac{2 \sum_{i: x_i \in [0.5; 1]} y_i}{\sum_{i: x_i \in [0.5; 1]} (x_i - 0.5)^2}.$$

В более сложном случае мы будем минимизировать функцию двух переменных (α и β): $f(\alpha, \beta) = \sum_{i: x_i \in [0.5; x_1^r]} (y_i - \alpha(x_i - 0.5))^2 + \sum_{i: x_i \in [x_1^r; 1]} (y_i - \beta(x_i - x_1^r) - \alpha(x_1^r - x_0^r))^2$.

Значение для x_1^r выберем равным $x_1^r = (x_{\max} - 0,5)/2$, где $x_{\max} = \max(x_i)$. Для того чтобы найти минимум такой функции, нам понадобится найти решение системы с частными производными:

$$\begin{cases} \frac{\partial f(\alpha, \beta)}{\partial \alpha} = 0, \\ \frac{\partial f(\alpha, \beta)}{\partial \beta} = 0. \end{cases}$$

Найдем производные и поделим каждое уравнение на -2 :

$$\begin{cases} \sum_{i: x_i \in [0, 5; x_1^r]} (y_i - \alpha(x_i - x_0^r))(x_i - x_0^r) + \\ + \sum_{i: x_i \in (x_1^r; 1]} (y_i - (\beta(x_i - x_1^r) + \alpha(x_1^r - x_0^r)))(x_1^r - x_0^r) = 0; \\ \sum_{i: x_i \in (x_1^r; 1]} (y_i - (\beta(x_i - x_1^r) + \alpha(x_1^r - x_0^r)))(x_i - x_1^r) = 0. \end{cases}$$

Из второго уравнения выразим β :

$$\beta = \frac{\sum_{i: x_i \in (x_1^r; 1]} (y_i - \alpha(x_1^r - x_0^r))(x_i - x_1^r)}{\sum_{i: x_i \in (x_1^r; 1]} (x_i - x_1^r)^2}.$$

Введем дополнительные переменные:

$$S_1 = \frac{\sum_{i: x_i \in (x_1^r; 1]} y_i (x_i - x_1^r)}{\sum_{i: x_i \in (x_1^r; 1]} (x_i - x_1^r)^2},$$

$$S_2 = \frac{\sum_{i: x_i \in (x_1^r; 1]} (x_1^r - x_0^r)(x_i - x_1^r)}{\sum_{i: x_i \in (x_1^r; 1]} (x_i - x_1^r)^2}.$$

В новых переменных $\beta = S_1 - \alpha S_2$. Подставляя полученное выражение в первое уравнение системы, получаем выражение для α :

$$\alpha = \frac{\sum_{i: x_i \in [0, 5; x_1^r]} y_i (x_i - x_0^r) + \sum_{i: x_i \in (x_1^r; 1]} (y_i - S_1(x_i - x_1^r))(x_1^r - x_0^r)}{\sum_{i: x_i \in [0, 5; x_1^r]} (x_i - x_0^r)^2 + \sum_{i: x_i \in (x_1^r; 1]} ((x_1^r - x_0^r) - S_2(x_i - x_1^r))(x_1^r - x_0^r)}.$$

Для случая трех оптимизируемых параметров α , β и x_1^r мы будем использовать полученные выше выражения для вычисления α и β , а также какой-либо метод одномерной численной оптимизации для x_1^r .

Метод, основанный на минимизации ожидаемых потерь, заключается в численной оптимизации упрощенной стратегии каким-либо методом многопараметрической оптимизации (например, градиентным методом). Значение ожидаемых потерь вычисляется при помощи рекурсивных формул из [4] по упрощенной стратегии. Этот метод требует больше вычислений: необходимо упрощать стратегию целиком (от $(N-2)$ до $3 \cdot (N-2)$ переменных, в зависимости от выбранного числа оптимизируемых параметров для каждого шага), вместо упрощения стратегий для от-

дельных шагов (несколько простых задач, от 1 до 3 переменных в каждой).

Обратим внимание на предложенный в [5] способ параллельной обработки данных с различными размерами групп. Стратегия для такой обработки будет похожа на описанную выше с той лишь разницей, что не для каждого элемента матрицы A будет определено значение. Значения порогов будут определены только для тех значений n_1 , n_2 , которые могут быть получены в процессе управления с учетом размеров групп.

Для подобной обработки также применимы все предложенные методы поиска упрощенных стратегий.

4. Результаты экспериментов

При помощи метода, основанного на минимизации суммы квадратов отклонений от точной стратегии, были получены упрощенные стратегии для управления с равными ($N=15$ и $N=30$) и различными ($N=30$; $k=13$; размеры групп: 1, 1, 2, 1, 2, 1, 1, 2, 1, 2, 2, 2, 3, 3, 3, 6) размерами групп. Проведено моделирование управления с использованием полученных стратегий методом Монте-Карло.

На рис.4 показаны потери для управления с различными размерами групп. Жирной линией показаны потери для точной стратегии, тонкими линиями — потери для упрощенных стратегий. Как видно из рисунка, потери для разных способов упрощения не очень сильно отличаются между собой (выделяется только способ, основанный на кодировании каждого графика стратегии при помощи одной точки).

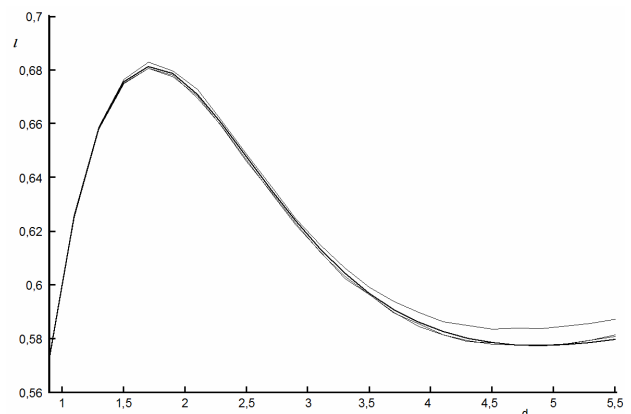


Рис.4. Сравнение потерь от применения точной и упрощенных стратегий

При нахождении и тестировании упрощенных стратегий получены следующие результаты: для того чтобы табулировать точную стратегию при $N=15$, потребуется хранить 42 точки (с учетом симметричности стратегий). Для табулирования упрощенной стратегии потребуется 13 или 26 точек (для кодирования одной и двумя точками соответственно). Разница между потерями для точной и упрощенных стратегий при моделировании методом Монте-Карло не превышала 0,8%.

Для того чтобы табулировать точную стратегию при $N=30$, потребуется хранить 197 точек (с

учетом симметричности стратегий). Для табулирования упрощенной стратегии потребуется 28 или 56 точек. Разница между потерями для точной и упрощенных стратегий не превышала при моделировании 0,9%.

Для управления с различными размерами групп точная стратегия потребует хранить 81 точку (с учетом симметричности стратегий), тогда как упрощенная стратегия потребует 13 или 26 точек. Разница между потерями для точной и упрощенных стратегий не превышала при моделировании 1,3%.

5. Заключение

Описанные способы получения упрощенных стратегий позволяют табулировать стратегии с использованием меньшего количества информации и при этом получать приемлемое ухудшение результатов. Описанные методы поиска упрощенных стратегий применимы к стратегиям параллельного управления с различными размерами групп, описанным в [5].

Автор выражает благодарность А.В.Колногорову за постановку задачи и помощь в подготовке статьи.

Работа выполнена при финансовой поддержке РФФИ, проект № 13-01-00334.

1. Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. М.: Наука, 1969. 316 с.
2. Срагович В.Г. Адаптивное управление. М.: Наука, 1981. 384 с.
3. Berry D.A., Fristedt B. Bandit Problems. L., N.Y.: Chapman and Hall. 1985. 275 p.
4. Колногоров А.В. Нахождение минимаксных стратегий и риска в случайной среде (задача о двуруком бандите) // Автоматика и телемеханика. 2011. №5. С.127-138.
5. Олейников А.О. Численная оптимизация параллельной обработки в стационарной случайной среде // Труды КарНЦ РАН. 2013. Т.1. С.73-74.

References

1. Tsetlin M.L. Issledovaniia po teorii avtomatov i modelirovaniu biologicheskikh sistem [The research on the automata theory and biological systems modelling]. Moscow, "Nauka" Publ., 1969. 316 p.
2. Sragovich V.G. Adaptivnoe upravlenie [Adaptive control]. Moscow. "Nauka" Publ., 1981. 384 p.
3. Berry D.A., Fristedt B. Bandit Problems. London, New York: Chapman and Hall. 1985, 275 p.
4. Kolnogorov A.V. Nakhozhdenie minimaksnykh strategii i riska v sluchainoi srede (zadacha o dvurukom bandite) [Finding minimax strategy and minimax risk in a random environment (the two-armed bandit problem)]. Avtomatika i telemekhanika – Automation and remote control, 2011, no. 5, pp. 127-138.
5. Oleinikov A.O. Chislennaia optimizatsiia parallel'noi obrabotki v statSIONARNOI sluchainoi srede [Numerical optimization of parallel processing in a stationary environment]. Trudy KarNTS RAN [Proceedings of the Karelian RC of the RAS]. 2013, vol. 1, pp. 73-74.