Труды Карельского научного центра РАН
№ 1. 2013. С. 73–78

УДК 519.865

ЧИСЛЕННАЯ ОПТИМИЗАЦИЯ ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ В СТАЦИОНАРНОЙ СЛУЧАЙНОЙ СРЕДЕ

А. О. Олейников

Новгородский государственный университет имени Ярослава Мудрого

В задаче о поведении в случайной среде (также известной как задача о двуруком бандите) предложена стратегия, которая применяет варианты к группам данных на конечном заданном числе этапов. Предложен алгоритм оптимизации размеров групп обрабатываемых данных. Представлены результаты численной оптимизации.

Ключевые слова: поведение в случайной среде, задача о двуруком бандите, робастное управление, параллельная обработка.

A. O. Oleynikov. NUMERICAL OPTIMIZATION OF PARALLEL PROCESSING IN A STATIONARY ENVIRONMENT

A stategy which applies variants to data groups in a bounded number of stages is considered for the problem of control in a stationary environment (also known as the Two-armed bandit problem). An algorithm is suggested for group size optimization. The results of the algorithm application are presented.

Key words: behavior in random environment, two-armed bandit problem, robust control, parallel processing.

Введение

Рассматривается задача о поведении в случайной среде [5], также известная как задача о двуруком бандите [4], [6] с нормально распределенными доходами на конечном числе шагов N [1]. Доходы при использовании различных вариантов имеют распределения с плотностями

$$f(x|m_\ell) = (2\pi)^{-1/2} \exp\left\{-(x-m_\ell)^2/2\right\},$$

где ℓ – номер выбранного варианта, а m_l – математическое ожидание этого варианта (в нашем случае $\ell = 1, 2$). Такой двурукий бандит описывается векторным параметром $\theta = (m_1, m_2)$.

Выбор варианта на следующем шаге осуществляется на основе информации о доходах на предыдущих шагах и описывается стратегией σ. Множество, состоящее из всех возможных стратегий, обозначим Σ. Целью управления является получение наибольшего дохода.

При известном параметре θ наилучшей стратегией является та, которая указывает всегда применять вариант, которому соответствует бóльшая из величин m_1, m_2. В таком случае полный ожидаемый доход равен $N(m_1 \vee m_2)$. Если же параметр неизвестен, то потери дохода вследствие неполноты информации описываются функцией

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

Здесь $E_{\sigma,\theta}$ – математическое ожидание по мере, порожденной стратегией σ и параметром θ. Множество допустимых значений параметра имеет вид $\Theta = \{(m_1, m_2) : |m_1 - m_2| \leqslant c\}$, где c – некоторая константа ($0 < c < \infty$) [1].

При использовании минимаксного подхода, цель управления состоит в минимизации максимальных ожидаемых потерь на множестве параметров Θ по множеству стратегий Σ, величина

$$R_N^M(\Theta) = \inf_{\Sigma} \sup_{\Theta} L_N(\sigma, \theta)$$

называется минимаксным риском, а соответствующая стратегия – минимаксной стратегией. Эта постановка рекомендована в работе [7].

Одним из возможных видоизменений стратегии является разбиение шагов на несколько групп, таким образом, чтобы каждый шаг принадлежал какой-либо группе и в одной группе могут оказаться только шаги, расположенные подряд. В дальнейшем будем называть шаги пакетами данных, а выбор одного из вариантов на каждом шаге – выбором варианта обработки для пакета.

В данной работе рассмотрен класс стратегий, применяющих одинаковый вариант для всех пакетов в группе и позволяющих вследствие этого вести их параллельную обработку. Для такого класса стратегий предложен алгоритм оптимизации размеров групп и представлены результаты численной оптимизации.

Стратегия параллельного управления

В [1] установлено, что для описанного случая применима основная теорема теории игр (минимаксные стратегия и риск совпадают с байесовскими на наихудшем априорном распределении). Также в этой работе установлено, что наихудшее априорное распределение является симметрическим и асимптотически однородным, и получено рекуррентное уравнение для вычисления соответствующих байесовских стратегии и риска.

Положим $m_1 = u + v$, $m_2 = u - v$, тогда $\theta = (u + v, u - v)$, $\Theta = \{\theta : |v| \leqslant c\}$. В новых переменных асимптотически наихудшая плотность распределения может быть выбрана в виде $\nu_a(u, v) = \varkappa_a(u)\rho(v)$, где $\varkappa_a(u)$ – постоянная плотность при $|u| \leqslant a$, $\rho(v) = \rho(-v)$ – симметрическая плотность и $a \to \infty$. Соответствующий байесовский риск равен [1]:

$$R_N^B(\nu_a(u,v)) \tag{1}$$
$$= \inf_{\Sigma} \iint_{\Theta} L_N(\sigma, (u+v, u-v))\nu_a(u,v)dudv.$$

Обозначим через n_1 и n_2 количество шагов, на которых применены первый и второй варианты соответственно, а через X_1, X_2 – полные доходы при их применении. При поиске байесовской стратегии на первых двух шагах оба варианта применяются по очереди, стратегию на следующих шагах можно найти, используя приведенные в [1] рекуррентные уравнения.

Далее рассмотрим стратегию, в которой используется разбиение шагов на $k+2$ группы и управление ведется с тем ограничением, что для всех пакетов в группе применяется один вариант [2], [3]. В начале стратегия указывает применять каждый вариант по M_0 раз (первые две группы), а затем осуществляет оптимальное управление с описанным выше ограничением. Размеры групп обозначим M_0 для первых двух и $M_1, M_2, \ldots, M_k$ для оставшихся (таким образом, первые две группы всегда имеют одинаковый размер). Считаем, что $2M_0 + M_1 + M_2 + \cdots + M_k = N$.

Такой вид управления позволяет производить параллельную обработку данных. Действительно, если в байесовской стратегии для выбора варианта для следующего пакета нужно знать результат обработки предыдущего, то для описанной стратегии возможно применение выбранного варианта ко всем пакетам в одной группе параллельно. В этом случае полное время управления равно времени обработки $k+1$ пакетов данных (поскольку первые две группы могут быть обработаны одновременно).

В работе [2] описывается похожая стратегия, но предписывающая разбивать обработку на группы равного размера. Разбиение же на группы разного размера вызвано желанием уменьшить максимальные потери при использовании параллельной стратегии. Как будет видно из результатов оптимизации, уменьшение размеров первых групп ведет к уменьшению максимальных рисков.

Обозначим через $R^B_{n_1,n_2}(Z)$ байесовский риск на последних $N - (n_1 + n_2)$ шагах относительно текущего апостериорного распределения, через $R^{B(\ell)}_{n_1,n_2}(Z)$ – аналогичный риск, вычисленный при условии, что сначала M_i раз выбирается ℓ-й вариант, а затем выполняется оптимальное управление, где $Z = X_1 n_2 - X_2 n_1$, $n_1 \geqslant M_0$, $n_2 \geqslant M_0$, $\ell = 1, 2$. Вычисления удобно выполнять для рисков $R_{n_1,n_2}(Z) = R^B_{n_1,n_2}(Z) p_{n_1,n_2}(Z)$, $R^{(\ell)}_{n_1,n_2}(Z) = R^{B(\ell)}_{n_1,n_2}(Z) p_{n_1,n_2}(Z)$, $\ell = 1, 2$, где

$p_{n_1,n_2}(Z)$ – плотность распределения Z при фиксированных n_1, n_2 [2].

При оптимизации использовалось следующее рекуррентное уравнение для вычисления байесовского риска, соответствующего стратегии параллельного управления [2], [3]:

$$R_{n_1,n_2}(\cdot) = \min(R^{(1)}_{n_1,n_2}(\cdot), R^{(2)}_{n_1,n_2}(\cdot)), \qquad (2)$$

где $R^{(1)}_{n_1,n_2}(Z) = R^{(2)}_{n_1,n_2}(Z) = 0$ при $n_1+n_2 = N$,

$$\begin{gathered} R^{(1)}_{n_1,n_2}(Z) = M_i g^{(1)}_{n_1,n_2}(Z) + n_2^{-1} \\ \times \int\limits_{-\infty}^{+\infty} R_{n_1+M_i,n_2}(Z+z) h_{n_1,M_i}\left(\frac{ZM_i - n_1 z}{n_2}\right) dz, \end{gathered} \qquad (3)$$

$$\begin{gathered} R^{(2)}_{n_1,n_2}(Z) = M_i g^{(2)}_{n_1,n_2}(Z) + n_1^{-1} \\ \times \int\limits_{-\infty}^{+\infty} R_{n_1,n_2+M_i}(Z+z) h_{n_2,M_i}\left(\frac{ZM_i - n_2 z}{n_1}\right) dz \end{gathered} \qquad (4)$$

при $n_1 + n_2 = 2M_0 + \cdots + M_{i-1}$, $n_1 \geqslant M_0$, $n_2 \geqslant M_0$. Здесь

$$\begin{gathered} g^{(\ell)}_{n_1,n_2}(Z) \\ = \int\limits_0^{\infty} 2v g_{n_1,n_2}(Z, (-1)^{\ell+1} v)\rho(v) dv, \quad \ell = 1, 2, \end{gathered} \qquad (5)$$

$$\begin{gathered} g_{n_1,n_2}(Z, v) = (2\pi n_1 n_2 (n_1+n_2))^{-1/2} \\ \times \exp\left(-\frac{(Z + 2v n_1 n_2)^2}{2 n_1 n_2 (n_1+n_2)}\right), \end{gathered} \qquad (6)$$

$$\begin{gathered} h_{n,M}(z) = \left(\frac{n+M}{2\pi n M}\right)^{1/2} \\ \times \exp\left(-\frac{z^2}{2nM(n+M)}\right). \end{gathered} \qquad (7)$$

При этом байесовский риск (1) вычисляется по формуле:

$$\begin{gathered} \lim_{a\to\infty} R^B_N(\nu_a(u,v)) \\ = 4M_0 \int_0^{\infty} v\rho(v) dv + \int_{-\infty}^{\infty} R_{M_0,M_0}(z) dz. \end{gathered} \qquad (8)$$

Запоминая для каждой тройки Z, n_1, n_2 минимальный риск R^ℓ, мы находим соответствующую байесовскую стратегию.

Для произвольной стратегии параллельного управления $\{\sigma_\ell(z, n_1, n_2)\}$ и плотности $\nu_a(u,v)$ вычислить потери можно, используя следующее рекуррентное уравнение [2]:

$$\begin{gathered} L_{n_1,n_2}(Z) = \sigma_1(Z, n_1, n_2) L^{(1)}_{n_1,n_2}(Z) \\ + \sigma_2(Z, n_1, n_2) L^{(2)}_{n_1,n_2}(Z), \end{gathered} \qquad (9)$$

где $L^{(1)}_{n_1,n_2}(Z) = L^{(2)}_{n_1,n_2}(Z) = 0$ при $n_1 + n_2 = N$, а функция $\sigma_\ell(Z, n_1, n_2)$ равна единице, если найденная байесовская стратегия предписывает на данном шаге при данных доходах выбирать ℓ-й вариант, и нулю – в противном случае.

$$\begin{gathered} L^{(1)}_{n_1,n_2}(Z) = M_i g^{(1)}_{n_1,n_2}(Z) + n_2^{-1} \\ \times \int\limits_{-\infty}^{+\infty} L_{n_1+M_i,n_2}(Z+z) h_{n_1,M_i}\left(\frac{ZM_i - n_1 z}{n_2}\right) dz, \end{gathered} \qquad (10)$$

$$\begin{gathered} L^{(2)}_{n_1,n_2}(Z) = M_i g^{(2)}_{n_1,n_2}(Z) + n_1^{-1} \\ \times \int\limits_{-\infty}^{+\infty} L_{n_1,n_2+M_i}(Z+z) h_{n_2,M_i}\left(\frac{ZM_i - n_2 z}{n_1}\right) dz \end{gathered} \qquad (11)$$

при $n_1 + n_2 = 2M_0 + \cdots + M_{i-1}$, $n_1 \geqslant M_0$, $n_2 \geqslant M_0$. Тогда функция потерь вычисляется по формуле

$$\begin{gathered} \lim_{a\to\infty} L_N(\nu_a(u,v)) \\ = 4M_0 \int_0^{\infty} v\rho(v) dv + \int_{-\infty}^{\infty} L_{M_0,M_0}(z) dz. \end{gathered} \qquad (12)$$

Оптимизация размеров групп данных

Для оптимизации размеров групп использовалась показанная выше методика нахождения минимаксного риска ((2) – (8)) и потерь ((9) – (12)) для заданного разбиения данных $M_0, M_1, \ldots, M_k$. Этот риск зависит от c и от разбиения $M_0, M_1, \ldots, M_k$. Задача состоит в поиске оптимального разбиения (для фиксированного c)

$$\begin{gathered} R(M^0_0, M^0_1, \ldots, M^0_k) \\ = \min_{M_0, M_1, \ldots, M_k} R(M_0, M_1, \ldots, M_k). \end{gathered}$$

Минимизация может выполняться численными методами, например, методом покоординатного спуска. Однако необходимо учитывать, что в данном случае переменные связаны между собой. Поэтому будем использовать следующую модификацию метода покоординатного спуска: будем поочередно находить оптимальные значения для каждого M_i, начиная с M_0. При этом все оставшиеся M_i будем считать равными (в случае, если оставшееся количество шагов не делится на количество оставшихся групп, необходимое количество последних групп увеличивается на единицу). Получившуюся подзадачу однопараметрической оптимизации будем решать прямым перебором. Полученное значение фиксируется, и оптимизация продолжается для следующей группы данных. Минимальное значение для M_i равняется единице, максимальное значение определяется в зависимости от i: для $M_0 \leqslant (N-k)/2$, для остальных

$$M_i \leqslant N - (k+1-i) - 2\dot{M}_0 - \sum_{j=1}^{j=i-1} M_j.$$

Прямой перебор возможен для одного параметра благодаря небольшим рассматриваемым значениям N, в дальнейшем он может быть заменен другим методом. Начало оптимизации с M_0 обусловлено тем, что данное значение сильно влияет на потери (так как каждый вариант, в том числе худший, должен быть применен M_0 раз в начале управления), а также тем, что на начальных этапах управления информированность о параметрах среды меняется быстрее, чем на последующих.

Для того, чтобы убедиться, что найденное разбиение действительно соответствует минимуму функции, выбранной в качестве критерия, проверяются разбиения, полученные из найденного путем изменения размеров двух групп на единицу. Если в процессе проверок находилось лучшее разбиение, оно в дальнейшем проверялось по тем же правилам.

Оптимизация проводилась по предложенному алгоритму для $N = 16, 30$ и $k = 6$. Для d были использованы значения от 0,9 до 5,5 включительно с шагом 0,2 (параметр d характеризует априорное распределение: считается, что плотность $\rho(v)$ сосредоточена в двух точках: $v = \pm dN^{-1/2}$). Критерием для сравнений разбиений являлся минимаксный риск (производилась минимизация данного критерия). В результате оптимизации были получены разбиения со следующими размерами: для

$N = 16$: $M_0^0 = 1$, $M_1^0 = 1$, $M_2^0 = 1$, $M_3^0 = 1$, $M_4^0 = 3$, $M_5^0 = 1$ и $M_6^0 = 7$. Для $N = 30$: $M_0^0 = 1$, $M_1^0 = 3$, $M_2^0 = 2$, $M_3^0 = 4$, $M_4^0 = 2$, $M_5^0 = 5$ и $M_6^0 = 12$.

Для получения результата при $N = 16$ потребовалось найти риски для 71 возможного разбиения, тогда как их общее количество для приведенного примера равняется:

$$\sum_{i=1}^{2i \leqslant N-k} C_{N-2i-1}^{k-1} = 1897.$$

В данном случае это количество равно количеству композиций числа N длины $k+2$, у которых первые два слагаемых равны.

Для $N = 30$ общее число вариантов равняется 210574. При использовании предложенного алгоритма потребовалось перебрать 283 варианта (включая проверки).

Таким образом, нам удалось значительно сократить перебор. Однако из-за применения метода покоординатного спуска есть риск нахождения локального минимума.

На рис. 1 показаны приведенные риски r для $N = 16$ ($r = \frac{R(M_0, M_1, \ldots, M_k)}{\sqrt{N}}$): линия 1 – для разбиения с одинаковыми размерами групп $M_i = 2$, линия 2 – для найденного оптимального разбиения, линия 3 – для такого же числа шагов для стратегии, в которой выбор можно менять на каждом шаге. Как видно на рисунке, потери для полученного разбиения меньше, чем для разбиения с одинаковыми размерами групп. Однако они выше, чем потери для стратегии с последовательным управлением, что является платой за возможность параллельной обработки.

На рис. 2 показаны приведенные риски и потери для параллельной (линии 1 и 2) и обычной (линии 3 и 4) стратегий ($N = 30$). На данном рисунке также видно, что применение параллельной стратегии увеличивает риски.

Заключение

Предложена стратегия параллельной обработки данных, предписывающая применять разные варианты обработки к группам последовательно поступающих пакетов. Минимаксный риск ищется с помощью рекуррентного уравнения как байесовский, соответствующий наихудшему априорному распределению, и зависит от разбиения групп обрабатываемых данных (для фиксированного c). Рассмотрена задача оптимизации разбиения данных для минимизации минимаксного риска.

В дальнейшем стратегия разбиения может быть развита добавлением ограничения на максимальный размер группы (например, для случая, когда максимальное количество одновременно обрабатываемых групп данных ограничено).

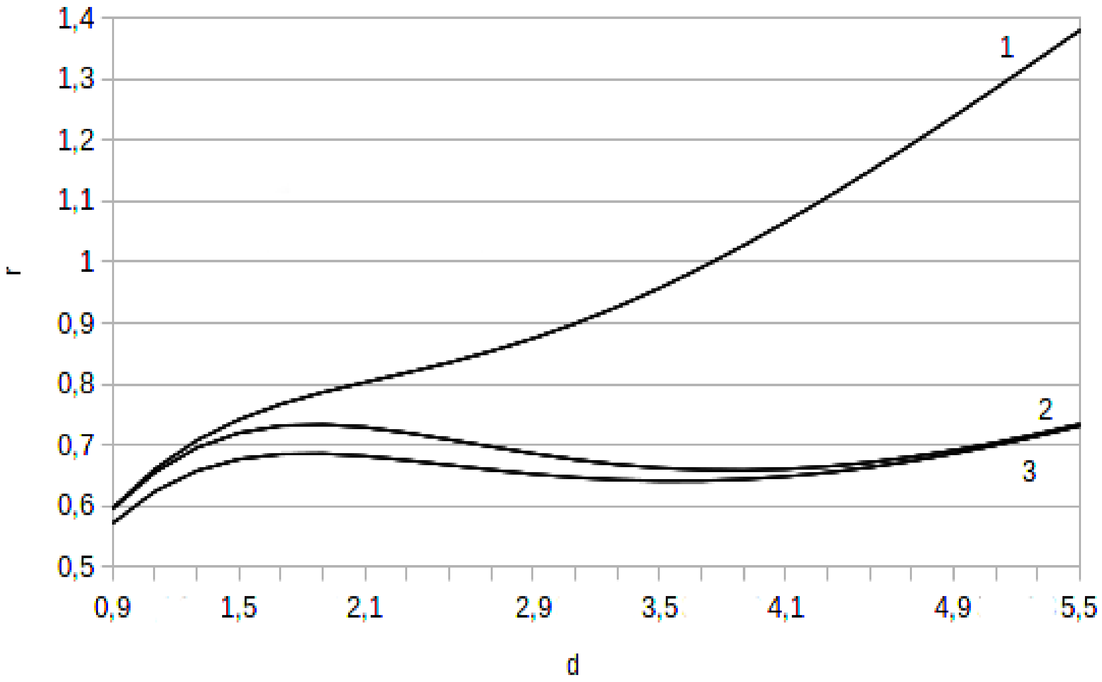

Рис. 1. Результаты численной оптимизации для 16 пакетов

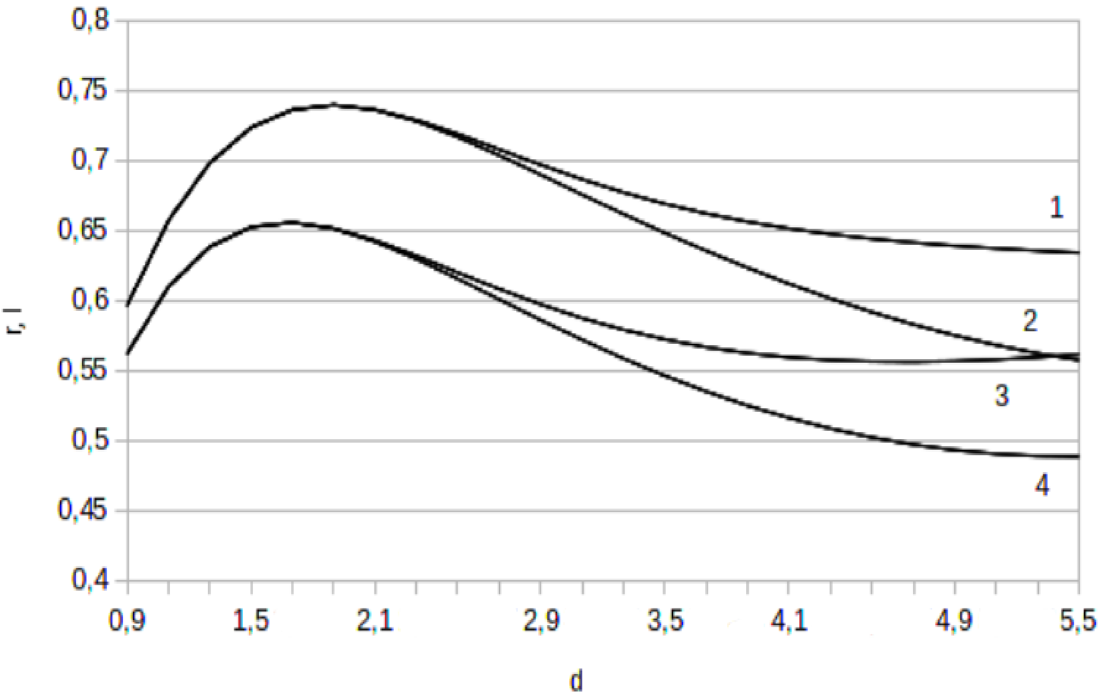

Рис. 2. Результаты численной оптимизации для 30 пакетов

Литература

1. *Колногоров А. В.* Нахождение минимаксных стратегии и риска в случайной среде (задача о двуруком бандите) // АиТ. 2011. № 5. С. 127–138.

2. *Колногоров А. В.* Робастное параллельное управление в случайной среде (задача о двуруком бандите) // АиТ. 2012. № 4. С. 114–130.

3. *Колногоров А. В., Олейников А. О.* Оптимизация параллельной многоэтапной обработки в случайной среде // Обозрение прикладной и промышленной математики. 2012. Т. 19, вып. 2. С. 210–211.

4. *Пресман Э. Л., Сонин И. М.* Последовательное управление по неполным данным. Байесовский подход. М.: Наука, 1982. 286 с.

5. *Цетлин М. Л.* Исследования по теории автоматов и моделированию биологических систем. М.: Наука, 1969. 316 с.

6. *Berry D. A., Fristedt B.* Bandit problems. London, New York: Chapman and Hall, 1985. 275 p.

7. *Robbins H.* Some aspects of the sequential design of experiments // Bulletin AMS. 1952. V. 58(5). P. 527—535.

СВЕДЕНИЯ ОБ АВТОРЕ:

Олейников Андрей Олегович
аспирант
Новгородский государственный университет имени Ярослава Мудрого
ул. Большая Санкт-Петербургская, 41, Великий Новгород, Новгородская область, Россия, 173000
эл. почта: duke.gh@yandex.ru
тел.: (8953) 901 64 23

Oleynikov, Andrey
Novgorod State University
41 Bol. Sankt-Peterburgskaya St., Velikiy Novgorod, Russia, 173000
e-mail: duke.gh@yandex.ru
tel.: (8953) 901 64 23