

А. О. Олейников (Великий Новгород, НовГУ). **Способы получения упрощенных стратегий для параллельной обработки в случайной среде.**

Рассматривается задача о двуруком бандите с нормально распределенными доходами на конечном отрезке времени длины N [1]. Предполагается, что текущие доходы ξ_n , $n = 1, \dots, N$ имеют разные мат. ожидания, но равные дисперсии, и зависят только от выбранного на текущем шаге варианта. Такой двурукий бандит характеризуется векторным параметром $\theta = (m_1, m_2)$. Цель управления состоит в максимизации полного ожидаемого дохода. Для этого используется стратегия σ . Функция потерь, обусловленная неполнотой информации: $L_N(\sigma, \theta) = E_{\sigma, \theta} \sum_{n=1}^N ((m_1 \vee m_2) - \xi_n)$. Байесовский и минимаксный риски определяются как

$$R_N^B(\lambda) = \min_{\{\sigma\}} \int_{\Theta} L_N(\sigma, \theta) \lambda(\theta) d\theta, \quad R_N^M(\lambda) = \min_{\{\sigma\}} \max_{\Theta} L_N(\sigma, \theta),$$

В работе [1] показано, что минимаксная стратегия и риск могут быть найдены как Байесовские при помощи рекуррентных формул. Стратегия имеет пороговый характер и может быть задана двумерной матрицей. Аналогичная матрица может быть получена для стратегий параллельной обработки (способ нахождения таких стратегий описан в [2]).

Также в [1] можно найти более наглядный способ описания таких стратегий. Для представления стратегии вводятся обозначения $t_1 = n_1/N$, $t_2 = n_2/N$, $t = t_1 + t_2$ и $T(t_1, t_2)$ — пороговое значение, определяющее, какой вариант необходимо выбрать на данном шаге. Для групп значений t равными t можно построить графики зависимости $T(t_1, t_2)$ от t_1/t . Каждый из этих графиков симметричен относительно $(0,5; 0)$.

Используя полученные в результате численных экспериментов стратегии, попытаемся описать способ получения приближенных стратегий, описываемых тремя точками для каждого t . Приближения могут быть осуществлены численными методами или аналитическими (с дополнительными допущениями). Для аналитического метода критерием оптимальности будет служить среднеквадратичное отклонение от исходного графика, для численного можно также использовать моделирование стратегии методом Монте-Карло.

Для аналитического метода стратегия для одного t состоит из точек $(0,5; 0)$, (x_1^r, y_1^r) , (x_2^r, y_2^r) . Возьмем в качестве x_2^r максимальный x приближаемой функции, а в качестве x_1^r какое-то значение из диапазона $(0,5; x_2^r)$.

$$y_1^r = (x_1^r - 0,5) \frac{\sum_{x_i \in [0,5; x_1^r]} (2x_i y_i - y_i)}{\sum_{x_i \in [0,5; x_1^r]} (\sqrt{2}x_i - \sqrt{0,5})^2}$$

$$y_2^r = y_1^r + (x_2^r - x_1^r) \frac{\sum_{x_i \in (x_1^r; 1]} (x_i y_i - y_i x_1^r - x y_1^r + y_1^r x_1^r)}{\sum_{x_i \in (x_1^r; 1]} (x_i - x_1^r)^2},$$

где (x_i, y_i) — точки приближаемой функции.

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

СПИСОК ЛИТЕРАТУРЫ

1. *Колмогоров А. В.* Нахождение минимаксных стратегии и риска в случайной среде (задаче о двуруком бандите). — Автомат. телемех., 2011, № 5, с. 127–138.
2. *Олейников А. О.* Численная оптимизация параллельной обработки в стационарной случайной среде. — Тр. КарНЦ РАН, 2013, № 1, с. 73–78.