ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 20 МАТЕМАТИКИ Выпуск 4
2013

А. Н. Лазутченко (Великий Новгород, НовГУ). **Использование двухпороговой стратегии управления в бинарной случайной среде.**

Случайная среда с бинарными доходами — это управляемый случайный процесс ξ_t, принимающий значения 0 и 1, интерпретируемые как текущие доходы и зависящие только от выбираемых в текущие моменты времени действий y_ℓ, т. е.

$$\mathbf{P}\{\xi_t = 1|y_t = \ell\} = p_\ell, \quad \mathbf{P}\{\xi_t = 0|y_t = \ell\} = q_\ell, \quad p_\ell + q_\ell = 1, \quad \ell = 1, 2; \quad t = 1, 2, \ldots, T.$$

Такая среда описывается векторным параметром $\theta = (p_1, p_2)$. Как следует из [1], в данной постановке задачи параметр фиксирован, но не известен лицу, осуществляющему управление. Рассматривается целевая функция потерь $L_T(\theta, \sigma)$, значениями которой являются потери за время моделирования, где θ определяет вероятности выигрыша на действиях, σ — используемая стратегия. Если параметр θ известен, то наилучшей стратегией является применять только то действие, которому соответствует большая из величин p_1, p_2, и максимальный полный доход в этом случае равен $\max\{p_1, p_2\}$. Если же θ неизвестен, то возникают потери вследствие неполноты информации о среде, равные $L_T(\theta, \sigma) = \max\{p_1, p_2\}T - E_{\sigma,\theta}(\sum_{t=1}^{T} \xi_t)$. Здесь $E_{\sigma,\theta}$ представляет собой математическое ожидание потерь полного дохода. Множество допустимых значений параметра θ имеет вид $\Theta = \{\theta\colon\ \epsilon \leqslant p_1 \leqslant 1 - \epsilon,\ \epsilon \leqslant p_2 \leqslant 1 - \epsilon,\ \epsilon > 0\}$. При использовании минимаксного подхода цель управления состоит в минимизации величины потерь полного дохода на множестве параметров Θ по множеству стратегий Σ. При этом минимаксный риск $R_T^M(\Theta) = \inf(\Sigma)\sup(\Theta)L_T(\theta, \sigma)$. Для управления предлагается использовать пороговую стратегию σ, которая применяет действия y_1 и y_2 по очереди, накапливая доходы X_1 и X_2 на действиях соответственно. На каждом шаге вычисляется абсолютная разность доходов $|X_1 - X_2|$ на действиях. Действия применяются до тех пор, пока эта величина не превысит порога $\alpha(DT)^{1/2}$, где T — полное время управления, α, D — пороговая константа и дисперсия соответственно $(D = p(1-p))$, или не истечет время управления. Если время управления не истекло, то действие, которому соответствует меньшая величина набранного дохода, исключается из рассмотрения, а d оставшееся время применяется только другое действие. Наибольшие потери полного дохода при достаточно больших T будут иметь место при $p_1 = 1/2$, $p_2 = p_1(1 - \beta T^{1/2})$, где $0 \leqslant \beta \leqslant T^{1/2}$. Ограничения на T накладываются, исходя из свойства инвариантности функции потерь.

На основе такой стратегии σ была разработана программа, которая рассчитывает оптимальные значения параметра α и β. Более конкретно: $\alpha = 0,58$, $\beta = 3,9$, $L_T = 0,375$ при $T = 10\,000$, $N = 1\,000\,000$.

Как показано в [1], вводя величину $S = \int_\beta L(\alpha, \beta) d\beta$, которая будет показывать суммарные потери на множестве значений β, можно рассмотреть суммарные минимальные потери для минимальных α для каждого β, а также суммарные потери для фиксированного β. Для значений β из промежутка [0, 50] вычисленные суммарные потери равны в первом случае 8,665, во втором 14,702. Очевидно, такая разница объясняется тем, что при больших значениях параметра β лучшее действие определяется достаточно быстро, но большое значение пороговой константы α не позволяет исключить из рассмотрения неоптимальное действие раньше, чем будет достигнут порог. Этот недостаток стратегии можно устранить, введя дополнительный порог, с помощью которого при больших β оптимальное действие будет определяться быстрее. В этом случае стратегия будет выглядеть следующим образом. Начиная с некоторого времени, текущий порог заменяется на исходный, после чего моделирование продолжается в обычном порядке. Для такой стратегии вычисленные суммарные потери дают значение 10,188. Как видно, введение второго порога позволяет в среднем снизить суммарные потери по сравнению с одиночным фиксированным порогом на 44%.

СПИСОК ЛИТЕРАТУРЫ

1. *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.