А. В. К о л н о г о р о в (Великий Новгород, НовГУ). **Нахождение минимаксных стратегии и риска в задаче о двуруком бандите с нормально распределенными доходами.**

Рассматривается задача о двуруком бандите на конечном отрезке времени длины N. Предполагается, что текущие доходы ξ_n, $n = 1, \ldots, N$, имеют нормальные распределения с плотностями $f(x|m_\ell) = (2\pi)^{-1/2} \exp\{-(x - m_\ell)^2/2\}$, где m_ℓ соответствует текущему выбранному варианту ($\ell = 1, 2$). Такой двурукий бандит характеризуется неизвестным векторным параметром $\theta = (m_1, m_2)$, для которого, однако, известно множество допустимых значений $\Theta = \{\theta\colon\ |m_1 - m_2| \leqslant c_1,\ |m_1 + m_2| \leqslant c_2\}$, где $0 < c_1 < \infty$, $0 < c_2 < \infty$, причем c_2 предполагается достаточно большим. Для управления применяется стратегия σ, которая может использовать всю известную предысторию (X_1, n_1, X_2, n_2) к текущему моменту времени $n = n_1 + n_2$, где n_1, n_2 характеризуют суммарные применения обоих вариантов, а X_1, X_2 — соответствующие полные доходы. Функция потерь, обусловленная неполнотой информации о процессе, имеет вид $L_N(\sigma, \theta) = E_{\sigma,\theta} \sum_{n=1}^{N}((m_1 \vee m_2) - \xi_n)$. Байесовский и минимаксный риски определяется как

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

им соответствуют байесовская и минимаксная стратегии. Из основной теоремы теории игр следует, что минимаксные стратегия и риск совпадают с байесовскими на наихудшем априорном распределении, соответствующем максимуму байесовского риска.

Как показано в [1], асимптотически наихудшая априорная плотность распределения обладает свойствами симметричности и асимптотической однородности. В переменных $u = (m_1 + m_2)/2$, $v = (m_1 - m_2)/2$, она может быть выбрана в виде $\lambda_a(u, v) = \kappa_a(u)\rho(v)$, где $\kappa_a(u)$ — постоянная плотность на отрезке $|u| \leqslant a$, $\rho(v)$ — симметрическая плотность (т. е. $\rho(-v) = \rho(v)$) на отрезке $|v| \leqslant c_1$ и a достаточно велико. Байесовская стратегия на первых двух шагах применяет варианты по очереди. Далее она определяется рекуррентно «с конца», т. е. надо вычислять риски $R_{n_1,n_2}(Z) = \min\{R^{(1)}_{n_1,n_2}(Z), R^{(2)}_{n_1,n_2}(Z)\}$, где $Z = X_1 n_2 - X_2 n_1$. Текущим оптимальным является ℓ-й вариант, если $R^{(\ell)}_{n_1,n_2}(Z)$ имеет меньшее значение. Здесь $R^{(1)}_{n_1,n_2}(Z) = R^{(2)}_{n_1,n_2}(Z) = 0$ при $n_1 + n_2 = N$ и далее

$$R^{(1)}_{n_1,n_2}(Z) = \int_0^\infty 2v g_{n_1,n_2}(Z, v)\rho(v)\, dv + \frac{1}{n_2}\int_{-\infty}^{+\infty} R_{n_1+1,n_2}(Z + z) h_{n_1}\left(\frac{Z - n_1 z}{n_2}\right) dz,$$

$$R^{(2)}_{n_1,n_2}(Z) = \int_0^\infty 2v g_{n_1,n_2}(Z, -v)\rho(v)\, dv + \frac{1}{n_1}\int_{-\infty}^{+\infty} R_{n_1,n_2+1}(Z + z) h_{n_2}\left(\frac{Z - n_2 z}{n_1}\right) dz,$$

при $n_1 + n_2 < N$, $n_1 \geqslant 1$, $n_2 \geqslant 1$. Функции $g_{n_1,n_2}(Z, v)$, $h_n(z)$ равны

$$g_{n_1,n_2}(Z, v) = \exp\left\{ -\frac{(Z + 2vn_1n_2)^2}{2n_1n_2(n_1 + n_2)} \right\}(2\pi n_1 n_2(n_1 + n_2))^{-1/2},$$

$$h_n(z) = \left(\frac{n+1}{2\pi n}\right)^{1/2} \exp\left\{ -\frac{z^2}{2n(n+1)} \right\}.$$

Тогда

$$\lim_{a\to\infty} R_N^B(\lambda_a(u, v)) = 4\int_0^\infty v\rho(v)\, dv + \int_{-\infty}^\infty R_{1,1}(z)\, dz.$$

По этим формулам численная оптимизация проводилась в предположении, что плотность $\rho(v)$ сосредоточена в двух точках $v = \pm dN^{-1/2}$ с вероятностями 0,5.

Тогда наихудшее априорное распределение соответствует максимуму приведенного байесовского риска $r_N(d) = N^{-1/2} R_N^B(\cdot)$. При $N = 30$, 50 максимумы $r_N(d)$ по $d = 0,9,\ 1,1,\ \ldots,\ 5,5$ оказались приблизительно равными 0,66, 0,65 соответственно при $d = 1,7$. Для найденных стратегий затем вычислялись приведенные потери $l_N(d) = N^{-1/2} L_N(\cdot)$, которые также имели указанные максимумы при $d = 1,7$.

СПИСОК ЛИТЕРАТУРЫ

1. *Колногоров А. В.* Свойства симметричности и асимптотической однородности «наихудшего» априорного распределения в задаче о двуруком бандите. — Обозрение прикл. и промышл. матем., 2009, т. 16, в. 2, с. 263–264.