Иерархическая модель конечной игры с природой

Краснова Мария Михайловна

Студентка

Факультет информатики ГГТУ, Орехово-Зуево, Россия

E-mail: mary.miyako@yandex.ru

В статье формализуется новая игровая модель принятия решений в условиях действия неконтролируемых (неопределенных) факторов. Принципиальное отличие от известных моделей [1] состоит в том, что «природа» реагирует на выбор лица, принимающего решения (ЛПР) своих стратегий, изменяя область возможных вероятностных распределений.

Рассматривается конечная игра с природой

$$\Gamma = \langle X, Y, \{P[x] \mid x \in X\}, \{f(x,p)\}\rangle.$$

Здесь набор $X = \{x_1, x_2, ..., x_m\}$ - совокупность альтернатив (стратегий) лица, принимающего решения (ЛПР). Набор

$$P[x_i] = \left\{ p^{(i)} \in R^n \mid (\forall_j \in \{1,2,...,n\})(b_j^{(i)} \le p_j^{(i)} \le c_j^{(i)}), \sum_{j=1}^{n} p_j^{(i)} = 1 \right\}$$

есть множество вероятностных распределений $p^{(i)} \in P[x_i]$ на множестве $Y = (y_1, y_2, ..., y_n)$, которые могут реализоваться в результате выбранного ЛПР решения $x_i \in X$. При этом выполнено очевидное включение $(\forall_j \in \{1,2,...,n\})[b_j^{(i)}, c_j^{(i)}] \subseteq [0,1]$, а также множество $P[x_i] \neq \oslash$ для любой стратегии $x_i \in X$.

Выигрыши ЛПР представлены матрицей $A = \| a_{ij} \|$, где выбор i- ой строки соответствуют возможным действиям (стратегиям) x_i ЛПР, а выбор столбца соответствует состояниям y_i природы. Функция выигрыша ЛПР определена как

$$f(x_i, p^{(i)}) = \sum_{j=1}^{n} p_j^{(i)} a_{ij}, x_i \in X, p^{(i)} \in P[x_i].$$

Перейдем к иерархической «процедуре» принятия решений в игре Γ . Первый ход делает ЛПР, именно он выбирает стратегию $x_i \in X$. Второй ход делает «природа», которая реализует информированную неопределенность $p^{(i)} \in P[x_i]$. Согласно подходу, в работе [2] стратегический риск по Вальду для ЛПР положим

$$R_V(x_i) = \max_{x_i \in X} \min_{p^{(i)} \in P[x_i]} f(x_i, p^{(i)}) - \min_{p^{(i)} \in P[x_i]} f(x_i, p^{(i)}),$$

стратегическое сожаление по Сэвиджу определим как

$$R_S(x_i) = \max_{x_i \in X} \varphi(x_i, p^{(i)}) - \min_{x_i \in X} \max_{p^{(i)} \in P[x_i]} \varphi(x_i, p^{(i)}),$$

где функция сожаления $\varphi(x_i, p)$ вычисляется согласно равенству

$$\varphi(x_i, p^{(i)}) = \max_{z_i \in X} f(z_i, p^{(i)}) - f(x_i, p^{(i)}).$$

В данной работе формализуем оптимальное поведение ЛПР в игре Γ , которое базируется на принципах Вальда и Сэвиджа из теории задач при неопределенности. Разумный игрок должен учитывать как возникающие риски при принятии решения, так и возможность получения большего выигрыша.

Определение 1. *Решение $x_U \in X$ в исходной задаче назовем U-оптимальным по рискам и сожалениям, если оно является минимальным для функции $F(x) = R_V^2(x) + R_S^2(x), x \in X$.*

Теорема 1. *В исходной задаче Γ существует U-оптимальное по рискам и сожалениям решение, при этом оно может быть найдено последовательным решением соответствующих задач математического программирования.*

Литература

1. Жуковский В. И., Кудрявцев К. Н., Смирнова Л. В. Гарантированные решения конфликтов и их приложения. М.: КРАСАНД, 2013.
2. Бардин А. Е., Житенева Ю. Н. Игры гарантий двух лиц // Межд. конф. КРОМШ - 2014. Спектральные и эволюционные задачи. Сборник тезисов. 2003. С. 80.