マルコフ決定過程に於けるリスク評価

リスク解析戦略研究センター
特任研究員　　影山 正幸

1　はじめに

Markov Decision Processes(以後, MDPs) の利得に対するリスク評価をファイナンスの分野でその有用性が指摘されている Value at Risk(VaR) [2] と Conditional Value at Risk(CVaR) を適応させて議論する. MDPs のリスクに関する研究は Sobel [6] によりはじまり, その後, Wu [7] により optimal policy の存在定理, 具体的に解をみつけるためのアルゴリズムが開発された. 一方, [3] により裾の重い分布に対して標準偏差よりも有効な VaR, CVaR を MDPs のリスク評価基準として用いそれぞれに対する optimal policy の存在定理と両者の関係式が与えられた. ここでは [3] に従って議論を行う.

2　Risk measure の利用

2.1　Multi stage VaR の存在

Definition 2.1 *与えられた policy* $\pi \in \Pi_0$*, 初期状態* $i \in S$*, threshold* α *に対して* $n-$*stage reward* R_n^π *に対する VaR とその optimal function* $\zeta^*_{n,\alpha}(i)$ *を次式で定義する.*

$$\begin{aligned} \zeta^\pi_{n,\alpha}(i) &:= -\sup\{\zeta | V^\pi_n(i,\zeta) \le \alpha\}, \forall i \in S, \pi \in \Pi_0, \alpha \in [0,1], n \ge 1. \\ \zeta^*_{n,\alpha}(i) &:= \inf_{\pi \in \Pi_0}\{\zeta^\pi_{n,\alpha}(i)\}, \forall i \in S, \pi \in \Pi_0, \alpha \in [0,1], n \ge 1. \end{aligned} \tag{2.1}$$

また, $V^*_n(i,x)$ も x の分布関数になることから同様にして VaR を次式で定義することができる.

$$x^*_{n,\alpha}(i) := -\sup\{x | V^*_n(i,x) \le \alpha\}, i \in S, \alpha \in [0,1], n \ge 1. \tag{2.2}$$

このとき, 次の定理が成り立つ [3].

Theorem 2.1 $\forall i \in S, \alpha \in S, n \ge 1$ *に対して* $x^*_{n,\alpha}(i) = \zeta^*_{n,\alpha}(i) = \inf_\pi \zeta^\pi_{n,\alpha}$ *が成り立ち, 次の関係式を満たす* $\alpha-$*optimal policy* $\hat{\pi}_\alpha \in \Pi_0$ *が存在する.*

$$\zeta^{\hat{\pi}_\alpha}_{n,\alpha}(i) = \zeta^*_{n,\alpha}(i).$$

この定理と $V^*_n(i,x)$ が最適方程式

$$\begin{aligned} V^*_0(i,x) &= I_{[0,+\infty)}(x), \forall i \in S, x \in \Re, \\ V^*_n &= \min_{a \in A(i)}\Big\{\sum_{j \in S} p^a_{ij} V^*_{n-1}(j,(x - r(i,a,j))/\beta)\Big\}, \forall i \in S, x \in \Re, n \ge 1 \end{aligned}$$

を満足することから Dynamic Programming により, 具体的に R^π_n に対する VaR を計算することができる. 次に凸性を満たすことから VaR よりも性質が良いとされる CVaR [5] を定義し同様の議論を行う.

2.2 Multi-stage CVaR

Definition 2.2 与えられた $\pi \in \Pi_0$ に対して $\alpha-CVaR$ と $y^*_{n,\alpha}(i)$ をそれぞれ次式で定義する.

$$
\begin{aligned}
\phi^{\pi}_{n,\alpha}(i) &:= \frac{1}{\alpha}\int_0^{\alpha} \zeta^{\pi}_{n,p}(i)dp, \\
y^*_{n,\alpha}(i) &:= \frac{1}{\alpha}\int_0^{\alpha} x^*_{n,p}(i)dp.
\end{aligned} \tag{2.3}
$$

このとき, 次の定理が成り立つ.

Theorem 2.2 もし, $A^*_k(i) \neq \emptyset, \forall i \in S, k = 1, \ldots, n$ ならば, 全ての $i \in S, \alpha \in [0,1], n \geq 1$ に対して $y^*_{n,\alpha}(i) = \phi^*_{n,\alpha}(i) := \inf_{\pi \in \Pi_0} \phi^{\pi}_{n,\alpha}(i)$ が成り立ち, 次式を満たす *policy* $\hat{\pi} \in \Pi_0$ が存在し最適性の原理を満たす.

$$\phi^{\hat{\pi}}_{n,\alpha}(i) = \phi^*_{n,\alpha}(i).$$

ただし, $A^*_k(i)$ は optimal action sets [7]. よって, CVaR も VaR 同様に具体的に計算することが可能.

謝辞

本研究は NEDO(新エネルギー・産業技術総合開発機構) の公募研究「化学物質の最適管理をめざすリスクトレードオフ解析手法の開発」の一部としておこなわれた研究である.

参考文献

[1] Andre Philipp Mundt, Dynamic risk management with Markov decision processes, Universitat Karlsruhe (2007)

[2] P.Artzner, F.Delbaen, JM.Eber and D.Heath, Coherent Measure of Risk, *Mathematical Finance*, 9:203-227 (1999)

[3] King Boda and Jerzy A.Filar, Time Consistent Dynamic Risk Measures, Mathematical Methods in Operations Research 2005, Special issue in honor of Arice Hordijk 1-19 (2005)

[4] J.Goto and Y.Takano, Newsvendor solutions via conditional value-at-risk minimization, *European Journal Operational Research*, **179**, 80-96 (2007)

[5] RT.Rockafellar and S.Uryasev, Conditional value-at-risk for general loss distributions, *Journal of Banking & Finance*, 26:1443-1471 (2002)

[6] M.J.Sobel, The variance of discounted Markov decision processes, *J. Appl. Probab.*, **19**, 794-802 (1982)

[7] C.Wu and Y.Lin, Minimizing risk models in Markov Decision Processes with Policies Depending on Target Values, *J. Math. Anal. Appl.*, **231**, 47-67 (1999)