数理解析研究所講究録
第 897 巻 1995 年 80-85

Fenchel duality の応用

新潟大学·理学部数学科　田中謙輔 (Kensuke Tanaka)
新潟大·大学院自然科学研究科　星野満博 (Mitsuhiro Hoshino)
新潟大·大学院自然科学研究科　黒岩大史 (Daishi Kuroiwa)

In this paper, stochastic control processes have been investigated as dynamic programming models with an infinite horizon. In many cases, it is our main purpose to seek for an optimal policy under the various conditions. However, optimal policy may not exist under weak conditions. Then, under such weak conditions, we introduce a modified form of the dynamic model, which we want to call as dual dynamic one, and show that optimal value of the model is equal to one of the dual model. Moreover, we show that there exists an optimal policy for the dual model.

Key words: dynamic programming, optimal policy, Fenchel inequality, and Fenchel's duality theorem

1 制御 D.P モデル の記号と構成

次のような簡単な D.P モデル $(S, A, B, H, p, r, \beta)$ について考察する. ただし,

(1) $S = \{1, 2, 3, \cdots, i, \cdots\}$, システムの状態空間.

(2) A, B, システムの制御空間で共に Banach space.

(3) $A(i)$, 各状態 $i \in S$ に対応する許容制御集合 $A(i) \subset A$.

(4) H_i, 各状態 $i \in S$ に対応する連続な線形写像 $H_i \in L(A, B)$.

(5) p, 各 $(i, H_i a) \in S \times B$ に対応する S 上の確率.

(6) $r(i, a)$, 有界な実数値損失関数 $r : S \times A \to R$.

(7) β, 割引因子 $0 \leq \beta < 1$.

この時, 制御政策 $\pi = (f_1, f_2, \cdots, f_t, \cdots)$, $f_t : S \to A$, のみについて考察し, 各制御 $f_t, t = 1, 2, 3, \cdots$, は現時点での状態のみに依存するマルコフ制御と呼ばれており, このような制御政策の全体を記号 Π で表す. 更に, 各初期状態 $i \in S$ と制御政策 $\pi \in \Pi$ に対応する総期待損失を

$$I(\pi)(i) = \sum_{t=1}^{\infty} \beta^{(t-1)} E_\pi[r(s_t, f_t(s_t))|s_1 = i]$$

で与えられる. この時, 確率過程 $\{s_t\}_{t=1,2,\ldots,}$ は Markov chain を構成している. ここで, 最適化問題を次のように表現する :

$$\text{(MP)} \qquad \text{minimize} \quad I(\pi)(x) \quad \text{subject to} \quad \pi \in \Pi.$$

Definition 1 この (MP) 問題で

$$\overline{I}(i) = \inf_{\pi \in \Pi} I(\pi)(i)$$

を満たす値 $\overline{I}(i)$ は初期状態 $i \in S$ に対応する最適損失値と呼ぶ

Definition 2 この (MP) 問題で, もし $\overline{I}(i) = I(\overline{\pi})(i)$ を満たす制御政策 $\overline{\pi} \in \Pi$ が存在すれば, この $\overline{\pi} \in \Pi$ を初期状態 $i \in S$ に対応する最適制御政策と呼ぶ. 更にこの $\overline{\pi} \in \Pi$ が全ての初期状態 $i \in S$ に対応する最適制御政策であるとき, 単に最適制御政策と呼ぶ.

2 補助定理と定理の証明

本論の主な補助定理と定理の証明を与える為に次の記号を用いる. この章を通して S 上の有界実数値関数の全体を $B(S)$ とおき, $u \in B(S)$ に対して次のような各記号を導入する：

1. $F(i,a,u) = r(i,a) + \beta \sum_{j \in S} u(j)p(j|i, H_i a) = r(i,a) + \beta G(H_i a, u)(i)$
2. $v(i) = \inf_{a \in A(i)} F(i,a,u)$
3. $r^*(i,p) = \sup_{a \in A(i)}[\langle a, p \rangle - r(i,a)]$, $p \in A^*$. ここで A^* は A の双対空間
4. $(\beta G(\cdot, u)(i))^*(q) = \beta \sup_{b \in B}[\langle \frac{q}{\beta}, b \rangle - G(b,u)(i)] = \beta G^*(\frac{q}{\beta}, u)(i)$, $q \in B^*$.ここで B^* は B の双対空間
5. $F^*(i,q,u) = r^*(i, -H_i^* q) + \beta G^*(\frac{q}{\beta}, u)(i)$, H_i^* は H_i の共役作用素で $H_i^* \in L(B^*, A^*)$
6. $v^*(i) = \inf_{q \in B^*} F^*(i,q,u)$

この時, $v(i), v^*(i)$ に Fenchel の不等式を適用して, 次の補助定理が示される.

Lemma 1 全ての $i \in S$ に対して,

$$v(i) + v^*(i) \geq 0.$$

Proof 各 $i \in S, u \in B(S)$ について，全ての $a \in A, q \in B^*$ に対して，Fenchel の不等式を適用して

$$\begin{aligned} F(i,a,u) + F^*(i,q,u) &= r(i,a) + \beta G(H_i a, u)(i) + r^*(i, -H_i^* q) + \beta G^*(\frac{q}{\beta}, u)(i) \\ &\geq \langle -H_i^* q, a\rangle + \beta \langle \frac{q}{\beta}, H_i a\rangle \\ &= 0 \end{aligned}$$

上の不等式で $a \in A$ と $q \in B^*$ について下限を取る事より結論の不等式が得られる． □

次に，各 $i \in S$ に対して，$\mathrm{dom}\, r^*(i,\cdot) = A^*(i), \mathrm{dom} G^*(\cdot, u)(i) = B^*(i)$ に対して，以下のような写像と凸錐を導入する：

$$\Psi_i : A^*(i) \times B^*(i) \to R \times A^*$$

$$\Psi_i(p,q) = \left(r^*(i,p) + \beta G^*(\frac{q}{\beta}, u)(i),\ p + H_i^* q \right)$$

$$Q = [0, \infty) \times \{\theta\} \subset R \times A^*$$

上の写像を用い，Fenchel の双対定理を示す証明の流れの一部を適用し，次の補助定理が得られる．

Lemma 2 全ての $i \in S,\ u \in B(S)$ に対して，$\mathrm{dom} G(\cdot, u)(i) = B(i) \subset B$ とおき，

$$\theta \in \mathrm{int}(H_i A(i) - B(i)),\ \theta \in H_i^* B^*(i) + A^*(i)$$

を満たすとする．ただし記号 int は集合の内点の全体．この時，次の式を満たす写像 $f^* : S \to B^*$ が存在する．

$$v^*(i) = r^*(i, -H_i^* f^*(i)) + \beta G^*(\frac{f^*(i)}{\beta}, u)(i)$$

Proof 各状態 $i \in S$ と $u \in B(S)$ に対して，結果の式を満たす B^* の対応する点が存在する事を示せば証明は終わる．そこで，$v^*(i)$ の定義から $v^*(i)$ に収束し，次の条件を満たす実数列 $v_n^*(i)$ が存在する：

$$v_n^*(i) = r^*(i, p_n) + \beta G^*(\frac{q_n}{\beta}, u)(i),$$

$$(v_n^*(i), p_n + H_i^* q_n) \in \Psi_i(A^*(i) \times B^*(i))$$

ただし，$r_n = p_n + H_i^* q_n$ とおくとき，全ての $n = 1, 2, \cdots,$ に対して，$r_n = \theta$ である．この補助定理の条件より，次の条件を満たす半径 $\varepsilon > 0$ の球 B_ε が存在する：

$$B_\varepsilon \subset \mathrm{int}(H_i A(i) - B(i))$$

ここで，任意の $z \in B$ に対して，$\frac{\varepsilon}{||z||}z = b - H_i a$ と書ける $a \in A(i), b \in B(i)$ が存在する．この事より，

$$\begin{aligned}
\frac{\varepsilon}{||z||}\langle z, q_n\rangle &= \langle b, q_n\rangle - \langle H_i a, q_n\rangle \\
&= \langle b, q_n\rangle + \langle a, p_n\rangle \\
&\leq r^*(i, p_n) + \beta G^*(\frac{q_n}{\beta}, u)(i) + r(i, a) + \beta G(b, u)(i) \\
&= v_n^*(i) + r(i, a) + \beta G(b, u)(i)
\end{aligned}$$

が成立する．この時に $v_n^*(i)$ は収束する事から，全ての $z \in B$ に対して

$$\sup_{n \geq 0}\langle z, q_n\rangle < \infty$$

だ成立する．よって，q_n の列は弱*コンパクトとなり，B^* のある点 q^* に弱*収束する部分列 $q_{n'}$ があり，同時に部分列 $p_{n'}$ は $p^* = -H_i^* q^*$ に弱*収束する．更に r^*, G^* は弱*下半連続であるから

$$\begin{aligned}
r^*(i, p^*) + \beta G^*(\frac{q^*}{\beta}, u)(i) &\leq \liminf_{n\to\infty} r^*(i, p_n) + \liminf_{n\to\infty} \beta G^*(\frac{q_n}{\beta}, u)(i) \\
&\leq \liminf_{n\to\infty}(r^*(i, p_n) + \beta G^*(\frac{q_n}{\beta}, u)(i)) \\
&= \lim_{n\to\infty} v_n^*(i) = v^*(i)
\end{aligned}$$

が成立する．これらの事から次の事が成立している

$$r^*(i, p^*) + \beta G^*(\frac{q^*}{\beta}, u)(i) = v^*(i),\ p^* + H_i^* q^* = \theta$$

かくて，各 $i \in S$ に対して $f^*(i) = q^*$ とおくことで $f^* : S \to B^*$ が得られ，証明は終わる． □

Lemma 3 *Lemma 2* と同じ条件の下で，全ての $i \in S$ に対して，

$$(-v(i), \theta) \in \Psi_i(A^*(i), B^*(i)) + Q$$

を満たしていると仮定する．この時，全ての $i \in S$ に対して，

$$v(i) + v^*(i) = 0.$$

Proof Lemma 2 とこの補助定理の条件より全ての $i \in S$ に対して，

$$\begin{aligned}
-v(i) &\geq v^*(i) \\
&= \inf_{q \in B^*} F^*(i, q, u) \\
&= F^*(i, q^*, u) \\
&= r^*(i, p^*) + \beta G^*(\frac{q^*}{\beta}, u)(i)
\end{aligned}$$

を満たし，更に $p^* + H_i^* q^* = \theta$ を満たす $p^* \in A^*, q^* \in B^*$ が存在する．亦，Lemma 1 より，$v(i) + v^*(i) \geq 0$ 即ち，$v^*(i) \geq -v(i)$ が成立する事より，補助定理の結論が得られ証明は終わる． □

Theorem 1 *D.P* モデルが次の条件を満たしている．全ての $i \in S$ に対して，

1. $\theta \in H_i^* B^*(i) + A^*(i)$, $\theta \in \mathrm{int}(H_i A(i) - B(i))$
2. $(-v_t(i), \theta) \in \Psi_i(A^*(i), B^*(i)) + Q$, $t = 1, 2, \cdots$,

ただし，$v_t, v_t^*, t = 1, 2, \cdots$, は次のように逐次的に作成される $B(S)$ の実関数である．

$$v_0 \equiv 0,\ v_t(i) = \inf_{a \in A(i)} F(i, a, v_{t-1})$$

$$v_0^* \equiv 0,\ v_t^*(i) = \inf_{q \in B^*} F^*(i, q, -v_{(t-1)}^*))$$

この時，各 $i \in S$ と各 $t = 1, 2, \cdots$, に対して，$f_t^* : S \to B^*$ で

$$v_t^*(i) = r^*(i, -H_i^* f_t^*(i)) + \beta G^*(\frac{f_t^*(i)}{\beta}, -v_{(t-1)}^*)(i)$$

を満たす $\pi^* = (f_1^*, f_2^*, \cdots, f_t^*, \cdots)$ が存在し，

$$\inf_{\pi \in \Pi} I(\pi)(i) = -I^*(\pi^*)(i)$$

が成立する．ただし，

$$I^*(\pi^*)(i) = \lim_{t \to \infty} v_t^*(i)$$

Proof 定理の条件から

$$\inf_{\pi \in \Pi} I(\pi)(i) = \lim_{t \to \infty} v_t(i)$$

が成立している．よって，上の補助定理の結果から各 $i \in S$ と各 $t = 1, 2, \cdots$, に対して，

$$v_t(i) = \inf_{a \in A(i)} F(i, a, v_{t-1}) = -v_t^*(i) = -F^*(i, f_t^*(i), -v_{t-1}^*)$$

を初期条件 $v_0 \equiv 0, v_0^* \equiv 0$ から，各段階 $t = 1, 2, \cdots$, に繰り返し適用して定理の結果が得られ，証明は終わる． □

References

[1] J.P.Aubin, Optima and Equilibria – An Introduction to Nonlinear Analysis, Springer-Verlag, New York, 1993.

[2] J.P.Aubin, Mathematical Methods of Game and Economic Theory, Revised Edition, North-Holland, Amsterdam, 1982.

[3] J.P.Aubin, & I.Ekeland, Applied Nonlinear Analysis, Wiley-Interscience, 1984.

[4] J.P.Aubin, & H.Frankowska, Set-Valued Analysis, Birkhäuser, Boston, 1990.

[5] D.P.Bertsekas and S.E.Shreve, Stochastic Optimal Control: The Discrete Time Case, Academic Press, New York, 1978.

[6] D.Blackwell, Discrete dynamic programming, Ann. Math. Statist. 33 (1962) 719-726.

[7] D.Blackwell, Discounted dynamic programming, Ann. Math. Statist. 36 (1965) 226-235.

[8] R.M.Dudley, Real Analysis and Probability, Wadsworth & Brooks, 1989.

[9] E.B.Dynkin and A.A.Yushkevich, Controlled Markov Processes, Springer-Verlag, Berlin, 1979.

[10] I.Ekeland, On the variational principle, J.Math.Anal.Appl., 47 (1974) 324-353.

[11] I.Ekeland, Nonconvex minimization problems, Bull. Amer. Math., 47 (1979) 443-474.

[12] K.Hinderer, Foundations of non-stationary dynamic programming with discrete time parameter, Lecture Notes on Operations Research and Mathematical Systems 33, Springer-Verlag, Berlin, 1970.

[13] K.Tanaka, On discounted dynamic programming with constraints, J. Math. Anal. Appl. 155 (1991) 264-277.