2.4. 結合型動的計画

様々な状況に対応しうる多段決定過程問題を定式化する。

- システムの評価は加法型のみとは限らない
- 結合律を満たす演算子にまで拡張
- より多様な価値観を反映可能
- さらに結合型評価の関数まで扱う
  （他に複合型評価なども）

2.4.1. 結合型評価:結合演算子による評価

システムは，各期における状態と決定（最終期は状態のみ）に応じて

$$r_1(x_1,u_1),\ r_2(x_2,u_2),\ \ldots,\ r_N(x_N,u_N),\ r_G(x_{N+1})$$

で評価される．そして，履歴 $(x_1,u_1,x_2,u_2,\ldots,x_N,u_N,x_{N+1})$ に対しては，

結合型評価：$r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_N(x_N,u_N)\circ r_G(x_{N+1})$

により評価される．なお，$\circ$ は結合演算子（結合律を満たす演算子）である．

$\circ=+$ のとき，**加法型評価**

$$r_1(x_1,u_1)+r_2(x_2,u_2)+\cdots+r_N(x_N,u_N)+r_G(x_{N+1})$$

$\circ=\times$ のとき，**乗法型評価**

$$r_1(x_1,u_1)\times r_2(x_2,u_2)\times\cdots\times r_N(x_N,u_N)\times r_G(x_{N+1})$$

$\circ=\wedge$ のとき，**最小型評価**（ただし $a\wedge b:=\min(a,b)$）

$$r_1(x_1,u_1)\wedge r_2(x_2,u_2)\wedge\cdots\wedge r_N(x_N,u_N)\wedge r_G(x_{N+1})$$

目的関数（期待値）

結合型評価の期待値

$$\begin{aligned}&E[r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_N(x_N,u_N)\circ r_G(x_{N+1})]\\ =&\sum\sum_{(x_2,\ldots,x_{N+1})\in X\times\cdots\times X}\cdots\sum\{[r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_N(x_N,u_N)\circ r_G(x_{N+1})]\\ &\times p_1(x_2|x_1,u_1)p_2(x_3|x_2,u_2)\cdots p_N(x_{N+1}|x_N,u_N)\}\end{aligned}$$

もしくは、結合型評価の関数の期待値

$$\begin{aligned}&E[g(r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_N(x_N,u_N)\circ r_G(x_{N+1}))]\\ =&\sum\sum_{(x_2,\ldots,x_{N+1})\in X\times\cdots\times X}\cdots\sum\{[g(r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_N(x_N,u_N)\circ r_G(x_{N+1}))]\\ &\times p_1(x_2|x_1,u_1)p_2(x_3|x_2,u_2)\cdots p_N(x_{N+1}|x_N,u_N)\}\end{aligned}$$

を目的関数と考える．

例2.4.1

加法型の目的関数

$$E[r_1(x_1,u_1)+r_2(x_2,u_2)+\cdots+r_N(x_N,u_N)+r_G(x_{N+1})]$$

確率型の目的関数　(T:定数)

$$\begin{aligned}&P(r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_G(x_{N+1})\ge T)\\ =&\sum\sum_{(x_2,\ldots,x_{N+1});r_1(x_1,u_1)\circ\cdots\circ r_G(x_{N+1})\ge T}\cdots\sum\{[r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_G(x_{N+1})]\\ &\times p_1(x_2|x_1,u_1)p_2(x_3|x_2,u_2)\cdots p_N(x_{N+1}|x_N,u_N)\}\\ =&E[C_{[T,\infty)}(r_1(x_1,u_1)\circ r_2(x_2,u_2)\circ\cdots\circ r_G(x_{N+1}))]\end{aligned}$$

ただし

$$C_{[T,\infty)}(x)=\begin{cases}1 & x\in[T,\infty)\\ 0 & x\notin[T,\infty)\end{cases}$$

乗法型評価について

通常の再帰式風に計算を実行してみると

$$\begin{aligned}\mathop{\mathrm{Max}}_{x,y\in[-2,1]}(x\times y)&=\mathop{\mathrm{Max}}_{x\in[-2,1]}(x\times\mathop{\mathrm{Max}}_{y\in[-2,1]}y)\\ &=\mathop{\mathrm{Max}}_{x\in[-2,1]}(x\times 1)\\ &=\mathop{\mathrm{Max}}_{x\in[-2,1]}(x)\\ &=1\end{aligned}$$

実際は

$$\mathop{\mathrm{Max}}_{x,y\in[-2,1]}(x\times y)=(-2)\times(-2)=4$$

2.4.2. 乗法型評価最大化問題

初期状態 $x_1\in X$ に対し，次の問題を考える

$$\begin{array}{ll}\text{Maximize} & E^\sigma[r_1(x_1,u_1)\times r_2(x_2,u_2)\times\cdots\times r_N(x_N,u_N)\times r_G(x_{N+1})]\\ \text{subject to} & \text{(i)}\ \ x_{n+1}\sim p(\cdot|x_n,u_n)\quad n=1,2,\ldots N\\ & \text{(ii)}\ \ \sigma=\{\sigma_1,\sigma_2,\ldots,\sigma_N\}\in\Sigma\end{array}$$

ただし，Σ は一般政策全体を表すものとする．

上記の問題における状態決定列は，各 σ に対し以下で与えられる:

$$\begin{array}{lll}\sigma_1(x_1)=u_1 & \to\ p(\cdot|x_1,u_1)\sim x_2 & \to\\ \sigma_2(x_1,x_2)=u_2 & \to\ p(\cdot|x_2,u_2)\sim x_3 & \to\\ \vdots & \quad\vdots & \\ \sigma_N(x_1,x_2,\ldots,x_N)=u_N & \to\ p(\cdot|x_N,u_N)\sim x_{N+1} & \end{array}$$

埋め込み法

不変埋没原理

問題を考える際，与問題を含むより大きな問題のクラスを考え，その中で解法を導く．そして最終的に必要な解を求める考え方．
（DP 自身の再帰的考え方もこの原理に基づく.）

⇩

与問題に対し，新たに実パラメータ $\lambda \in \mathbf{R}$ を加えた次の問題を考える：

埋め込み問題

Maximize $E^{\sigma}[\lambda \times r_1(x_1,u_1) \times r_2(x_2,u_2) \times \cdots \times r_N(x_N,u_N) \times r_G(x_{N+1})]$
subject to (i) $x_{n+1} \sim p(\cdot|x_n,u_n) \quad n=1,2,\ldots N$
(ii) $\sigma = \{\sigma_1,\sigma_2,\ldots,\sigma_N\} \in \Sigma$

$\lambda = 1$ のとき，埋め込み問題は与問題と等価になる.

埋め込み問題に対する部分問題群

部分問題群

$$V^{N+1}(x_{N+1},\lambda) = \lambda \times r_G(x_{N+1}), \quad x_{N+1} \in X,\ \lambda \in \mathbf{R}$$

$$V^{n}(x_n,\lambda) = \underset{\sigma\in\Sigma^n}{\mathrm{Max}}\, E^{\sigma}[\ \lambda \times r_n(x_n,u_n) \times \cdots \times r_G(x_{N+1})], \quad x_n \in X,\ \lambda \in \mathbf{R}$$
$$n = 1,2,\ldots,N$$

ここで $V^1(x_1,1)$ が与問題の最適値を与えることは容易にわかる.

拡大状態空間

元の問題の状態空間 X に対し，パラメータの空間を付加した $X \times \mathbf{R}$ を拡大状態空間と呼ぶ.

拡大状態空間における状態決定列

与問題に対する初期状態 x_1 とパラメータ λ_1 に対し

$$u_1 = \pi_1(x_1,\lambda_1) \ \rightarrow \begin{cases} x_2 \sim p(\cdot\,|\,x_1,u_1) \\ \lambda_2 = \lambda_1 r_1(x_1,u_1) \end{cases}$$

$$\rightarrow\ u_2 = \pi_2(x_2,\lambda_2) \ \rightarrow \begin{cases} x_3 \sim p(\cdot\,|\,x_2,u_2) \\ \lambda_3 = \lambda_2 r_2(x_2,u_2) \end{cases}$$

$$\rightarrow\ \cdots$$

$$\rightarrow\ u_N = \pi_N(x_N,\lambda_N) \ \rightarrow \begin{cases} x_{N+1} \sim p(\cdot\,|\,x_N,u_N) \\ \lambda_{N+1} = \lambda_N r_N(x_N,u_N) \end{cases}$$

ただし，$\pi = \{\pi_1,\pi_2,\ldots,\pi_N\}$ は拡大状態空間上でのマルコフ政策 とする

定理2.4.1 （再帰式）

埋め込み問題に対する部分問題群の値関数 $V^n(x,\lambda)$ $(n=1,2,\ldots,N+1)$ に対し、次の再帰式が成り立つ

$$V^{N+1}(x,\lambda) = \lambda\, r_G(x), \quad x \in X,\ \lambda \in \mathbf{R}$$

$$V^{n}(x,\lambda) = \underset{u\in U}{\mathrm{Max}} \left[\sum_{y\in X} V^{n+1}(y,\lambda\, r_n(x,u))p(y|x,u) \right], \quad x \in X,\ \lambda \in \mathbf{R}$$
$$n = 1,2,\ldots,N$$

埋め込み問題に対し，この再帰式から導かれる拡大状態空間上での最適マルコフ政策を $\pi^* = \{\pi_1^*,\pi_2^*,\ldots,\pi_N^*\}$ であらわすとき，与問題（乗法型評価最大化問題）の**最適一般政策**

$$\sigma^* = \{\sigma_1^*,\sigma_2^*,\ldots,\sigma_N^*\}$$

は次のように構成される：

（拡大状態空間上ではマルコフ政策のなかに最適政策が存在する!）

最適一般政策の導出

$\sigma_1^*(x_1) = \pi_1^*(x_1,\lambda_1),$
$\quad \lambda_1 = 1,\ x_1 \in X$

$\sigma_2^*(x_1,x_2) = \pi_2^*(x_2,\lambda_2),$
$\quad \lambda_2 = \lambda_1\, r_1(x_1,\ \sigma_1^*(x_1)),\ x_1,x_2 \in X$

$\sigma_3^*(x_1,x_2,x_3) = \pi_3^*(x_3,\lambda_3),$
$\quad \lambda_3 = \lambda_2\, r_2(x_2,\ \sigma_2^*(x_1,x_2)),$
$\quad x_1,x_2,x_3 \in X$

$\vdots$

$\sigma_N^*(x_1,x_2,\ldots,x_N) = \pi_N^*(x_N,\lambda_N),$
$\quad \lambda_N = \lambda_{N-1}\, r_{N-1}(x_{N-1},\ \sigma_{N-1}^*(x_1,x_2,\ldots,x^{N-1})),$
$\quad x_1,x_2,\ldots,\ x_N \in X$

例題2.4.1 （確定システム、乗法型評価）

2 期間 - 3 状態 - 2 決定 問題:

Maximize $r_1(u_1)r_2(u_2)r_G(x_3)$
subject to (i) $x_{n+1} = f(x_n,u_n) \quad n=1,2$
(ii) $u_1,u_2 \in U$

ただし，データは以下のとおり：

$r_G(s_1) = 0.3, \quad r_G(s_2) = -1.0, \quad r_G(s_3) = 0.8$

$r_2(a_1) = -0.8 \quad r_2(a_2) = 0.6$

$r_1(a_1) = 0.5, \quad r_1(a_2) = -0.9$

$f(x,u)$

$x \setminus u$	a_1	a_2
s_1	s_2	s_3
s_2	s_1	s_2
s_3	s_3	s_1

$$X = \{s_1,\ s_2,\ s_3\}, \quad U_1(x) = U_2(x) = U = \{a_1,\ a_2\}$$
$$f_1 = f_2 = f, \quad r_n(x,u) = r_n(u)$$

数値例の計算（埋め込み過程）(1/4)

定理 2.3 より，次の再帰式を計算すればよい．

$$V^3(x,\lambda) = \lambda \times r_G(x) \quad x \in X,\ \lambda \in \mathbf{R}$$

$$V^2(x,\lambda) = \operatorname*{Max}_{u \in U} V^3(f(x,u), \lambda \times r_2(u)) \quad x \in X,\ \lambda \in \mathbf{R}$$

$$V^1(x,\lambda) = \operatorname*{Max}_{u \in U} V^2(f(x,u), \lambda \times r_1(u)) \quad x \in X,\ \lambda \in \mathbf{R}$$

まず，$V^3(x,\lambda) = \lambda r_G(x)$ より

$$V^3(s_1,\lambda) = 0.3\lambda, \qquad V^3(s_2,\lambda) = -\lambda, \qquad V^3(s_3,\lambda) = 0.8\lambda$$

次に，$V^2(x,\lambda)$ を計算する

$$\begin{aligned} V^2(s_1,\lambda) &= V^3(f(s_1,a_1), \lambda \times r_2(a_1)) \vee V^3(f(s_1,a_2), \lambda \times r_2(a_2)) \\ &= V^3(s_2, \lambda \times (-0.8)) \vee V^3(s_3, \lambda \times 0.6)] \\ &= 0.8\lambda \vee 0.48\lambda = \begin{cases} 0.8\lambda & \lambda \geq 0 \\ 0.48\lambda & \lambda < 0 \end{cases} \qquad \pi_2^*(s_1,\lambda) = \begin{cases} a_1 & \lambda \geq 0 \\ a_2 & \lambda < 0 \end{cases} \end{aligned}$$

数値例の計算（埋め込み過程）(2/4)

$$\begin{aligned} V^2(s_2,\lambda) &= V^3(f(s_2,a_1), \lambda \times r_2(a_1)) \vee V^3(f(s_2,a_2), \lambda \times r_2(a_2)) \\ &= V^3(s_1, \lambda \times (-0.8)) \vee V^3(s_2, \lambda \times 0.6) \\ &= (-0.24\lambda) \vee (-0.6\lambda) = \begin{cases} -0.24\lambda & \lambda \geq 0 \\ -0.6\lambda & \lambda < 0 \end{cases} \qquad \pi_2^*(s_2,\lambda) = \begin{cases} a_1 & \lambda \geq 0 \\ a_2 & \lambda < 0 \end{cases} \end{aligned}$$

$$\begin{aligned} V^2(s_3,\lambda) &= V^3(f(s_3,a_1), \lambda \times r_2(a_1)) \vee V^3(f(s_3,a_2), \lambda \times r_2(a_2)) \\ &= V^3(s_3, \lambda \times (-0.8)) \vee V^3(s_1, \lambda \times 0.6) \\ &= (-0.64\lambda) \vee 0.18\lambda = \begin{cases} 0.18\lambda & \lambda \geq 0 \\ -0.64\lambda & \lambda < 0 \end{cases} \qquad \pi_2^*(s_3,\lambda) = \begin{cases} a_2 & \lambda \geq 0 \\ a_1 & \lambda < 0 \end{cases} \end{aligned}$$

最後に，$V^1(x,\lambda)$ を計算する

$$\begin{aligned} V^1(s_1,\lambda) &= V^2(f(s_1,a_1), \lambda \times r_1(a_1)) \vee V^2(f(s_1,a_2), \lambda \times r_1(a_2)) \\ &= V^2(s_2, \lambda \times 0.5) \vee V^2(s_3, \lambda \times (-0.9))] \\ &= \begin{cases} (-0.12\lambda) \vee 0.576\lambda & \lambda \geq 0 \\ (-0.3\lambda) \vee (-0.162\lambda) & \lambda < 0 \end{cases} = \begin{cases} 0.576\lambda & \lambda \geq 0 \\ -0.3\lambda & \lambda < 0 \end{cases} \end{aligned}$$

$$\pi_1^*(s_1,\lambda) = \begin{cases} a_2 & \lambda \geq 0 \\ a_1 & \lambda < 0 \end{cases}$$

2.4.3. 結合型評価最大化問題

初期状態 $x_1 \in X$ に対し，次の問題を考える

$$\begin{aligned} &\text{Maximize } E^\sigma[r_1(x_1,u_1) \circ r_2(x_2,u_2) \circ \cdots \circ r_N(x_N,u_N) \circ r_G(x_{N+1})] \\ &\text{subject to (i) } x_{n+1} \sim p(\cdot\,|x_n,u_n) \quad n = 1,2,\ldots N \\ &\qquad\qquad\ \text{(ii) } \sigma = \{\sigma_1,\sigma_2,\ldots,\sigma_N\} \in \Sigma \end{aligned}$$

確率システム上での一般の結合型評価をもつ問題は，埋め込み過程の考え方により再帰的な解法を導くことができる．

再帰式（埋め込み過程）

$$V^{N+1}(x,\lambda) = \lambda \circ r_G(x), \quad x \in X,\ \lambda \in \mathbf{R}$$

$$V^n(x,\lambda) = \operatorname*{Max}_{u \in U}\left[\sum_{y \in X} V^{n+1}(y, \lambda \circ r_n(x,u)) p(y|x,u)\right], \quad x \in X,\ \lambda \in \mathbf{R}$$

$$n = 1,2,\ldots,N$$

例2.4.2（最小型評価最大化問題）

$$\begin{aligned} &\text{Maximize } E^\sigma[r_1(x_1,u_1) \wedge r_2(x_2,u_2) \wedge \cdots \wedge r_N(x_N,u_N) \wedge r_G(x_{N+1})] \\ &\text{subject to (i) } x_{n+1} \sim p(\cdot\,|x_n,u_n) \quad n = 1,2,\ldots N \\ &\qquad\qquad\ \text{(ii) } \sigma = \{\sigma_1,\sigma_2,\ldots,\sigma_N\} \in \Sigma \end{aligned}$$

再帰式（埋め込み過程）

$$V^{N+1}(x,\lambda) = \lambda \wedge r_G(x), \quad x \in X,\ \lambda \in \mathbf{R}$$

$$V^n(x,\lambda) = \operatorname*{Max}_{u \in U}\left[\sum_{y \in X} V^{n+1}(y, \lambda \wedge r_n(x,u)) p(y|x,u)\right], \quad x \in X,\ \lambda \in \mathbf{R}$$

$$n = 1,2,\ldots,N$$

問題2.4.1

次の問題に対し埋め込み法による再帰式を用いて解（最適値と最適政策）を求めよ．ただし，$X = \{s_1,s_2,s_3\}$，$U = \{a_1,a_2\}$ とし，初期状態は $x_1 = s_1$ に固定する．

$$\begin{aligned} &\text{Maximize } \quad E^\sigma[\,r_1(u_1) \wedge r_2(u_2) \wedge r_G(x_3)\,] \\ &\text{subject to } \quad \text{(i) } x_{n+1} \sim p(\cdot\,|x_n,u_n) \quad n = 1,2 \\ &\qquad\qquad\qquad \text{(ii) } \sigma = \{\sigma_1,\sigma_2\} \in \Sigma \end{aligned}$$

$$r_G(s_1) = 0.8 \qquad r_G(s_2) = 0.5 \qquad r_G(s_3) = 1.0$$

$$r_2(a_1) = 0.9 \quad r_2(a_2) = 0.7\ ; \qquad r_1(a_1) = 0.8 \quad r_1(a_2) = 1.0$$

$p(x_{t+1}\,|\,x_t,a_1),\ t = 1,2$

$x_t \backslash x_{t+1}$	s_1	s_2	s_3
s_1	0.2	0.8	0.0
s_2	0.3	0.5	0.2
s_3	0.6	0.0	0.4

$p(x_{t+1}\,|\,x_t,a_2),\ t = 1,2$

$x_t \backslash x_{t+1}$	s_1	s_2	s_3
s_1	0.0	0.7	0.3
s_2	0.6	0.0	0.4
s_3	0.1	0.2	0.7

2.4.4. 複合評価問題

· 2 つの 2 項演算子 $\circ, \diamond$ はそれぞれ左単位元 $\tilde{\lambda}, \tilde{\mu}$ をもつものとする
· 各期において 2 種類の利得関数 r_n, R_n を考える
· 簡略化のため次の省略形を用いる：

$$r_n := r_n(x_n,u_n), \quad R_n := R_n(x_n,u_n), \quad n = 1,2,\ldots,N$$

$$r_{N+1} := r_G(x_{N+1}), \quad R_{N+1} := R_G(x_{N+1})$$

· 複合関数 $h : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ を考える

複合型評価最大化問題

$$\text{Maximize } E^\sigma[h(r_1 \circ r_2 \circ \cdots \circ r_N \circ r_{N+1},\ R_1 \diamond R_2 \diamond \cdots \diamond R_N \diamond R_{N+1})]$$

$$\begin{aligned} &\text{subject to (i) } x_{n+1} \sim p(\cdot\,|x_n,u_n) \quad n = 1,2,\ldots N \\ &\qquad\qquad\ \text{(ii) } \sigma = \{\sigma_1,\sigma_2,\ldots,\sigma_N\} \in \Sigma \end{aligned}$$

埋め込み法

Maximize $E^{\sigma}[h(\lambda \circ r_1 \circ r_2 \circ \cdots \circ r_N \circ r_{N+1},\ \mu \diamond R_1 \diamond R_2 \diamond \cdots \diamond R_N \diamond R_{N+1})]$

subject to (i) $x_{n+1} \sim p(\cdot|x_n, u_n) \quad n = 1, 2, \ldots N$
(ii) $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_N\} \in \Sigma$

ここで，$\lambda = \tilde{\lambda}, \mu = \tilde{\mu}$ とおくと元の問題と等価になる．

部分問題群

$$V^{N+1}(x_{N+1}, \lambda, \mu) = h(\lambda \circ r_G(x_{N+1}),\ \mu \diamond R_G(x_{N+1})) \quad x_{N+1} \in X$$

$$V^n(x_n, \lambda, \mu) = \underset{\sigma \in \Sigma^n}{\mathrm{Max}}\, E^{\sigma}[h(\lambda \circ r_n(x_n, u_n) \circ \cdots \circ r_G(x_{N+1}),\ \mu \diamond R_n(x_n, u_n) \diamond \cdots \diamond R_G(x_{N+1}))] \quad x_n \in X,\ n = 1, 2, \ldots, N$$

ここで $V^1(x_1, \tilde{\lambda}, \tilde{\mu})$ が与問題の最適値を与えることは容易にわかる．

定理2.4.2（再帰式，複合評価）

$$V^{N+1}(x, \lambda, \mu) = h(\lambda \circ r_G(x),\ \mu \diamond R_G(x)) \qquad x \in X,\ \lambda \in \mathbf{R},\ \mu \in \mathbf{R}$$

$$V^n(x, \lambda, \mu) = \underset{u \in U}{\mathrm{Max}} \sum_{y \in X} V^{n+1}(y,\ \lambda \circ r_n(x, u),\ \mu \diamond R_n(x, u))p(y\,|\,x, u)$$
$$x \in X,\ \lambda \in \mathbf{R},\ \mu \in \mathbf{R},\quad n = 1, 2, \cdots, N$$

この再帰式から導かれる拡大状態空間（ここでは $X \times \mathbf{R} \times \mathbf{R}$）上での最適マルコフ政策を $\pi^* = \{\pi_1^*, \pi_2^*, \ldots, \pi_N^*\}$ であらわすとき，与問題（複合型評価最大化問題）の**最適一般政策**

$$\sigma^* = \{\sigma_1^*, \sigma_2^*, \ldots, \sigma_N^*\}$$

は次のように構成される：

最適一般政策の導出

$\sigma_1^*(x_1) = \pi_1^*(x_1, \lambda_1, \mu_1)$,
$\quad \lambda_1 = \tilde{\lambda},\ \mu_1 = \tilde{\mu},\ x_1 \in X$

$\sigma_2^*(x_1, x_2) = \pi_2^*(x_2, \lambda_2, \mu_2)$,
$\quad \lambda_2 = \lambda_1 \circ r_1(x_1,\ \sigma_1^*(x_1)),\ \mu_2 = \mu_1 \diamond R_1(x_1,\ \sigma_1^*(x_1)),\ x_1, x_2 \in X$

$\sigma_3^*(x_1, x_2, x_3) = \pi_3^*(x_3, \lambda_3, \mu_3)$,
$\quad \lambda_3 = \lambda_2 \circ r_2(x_2,\ \sigma_2^*(x_1, x_2)),\ \mu_3 = \mu_2 \diamond R_2(x_2,\ \sigma_2^*(x_1, x_2))$,
$\quad x_1, x_2, x_3 \in X$

$\vdots$

$\sigma_N^*(x_1, x_2, \ldots, x_N) = \pi_N^*(x_N, \lambda_N, \mu_N)$,
$\quad \lambda_N = \lambda_{N-1} \circ r_{N-1}(x_{N-1},\ \sigma_{N-1}^*(x_1, x_2, \ldots, x^{N-1}))$,
$\quad \mu_N = \mu_{N-1} \diamond R_{N-1}(x_{N-1},\ \sigma_{N-1}^*(x_1, x_2, \ldots, x^{N-1}))$,
$\quad x_1, x_2, \ldots,\ x_N \in X$

例2.4.3（範囲型評価最小化問題）

minimize $E^{\sigma}[(r_1 \vee r_2 \vee \cdots \vee r_N \vee r_{N+1}) - (r_1 \wedge r_2 \wedge \cdots \wedge r_N \wedge r_{N+1})]$

subject to (i) $x_{n+1} \sim p(\cdot|x_n, u_n) \quad n = 1, 2, \ldots N$
(ii) $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_N\} \in \Sigma$

これは一般の複合型評価問題において

$$R_n = r_n,\quad h(\xi, \eta) = \xi - \eta,\quad \circ = \vee,\quad \diamond = \wedge$$

とおいた場合にあたる．

範囲型評価問題に対する部分問題群は

$$V^{N+1}(x_{N+1}, \lambda, \mu) = (\lambda \vee r_{N+1}) - (\mu \wedge r_{N+1})$$

$$V^n(x_n, \lambda, \mu) = \min_{\sigma \in \Sigma^n}\ E^{\sigma}\,[(\lambda \vee r_n \vee \cdots \vee r_{N+1}) - (\mu \wedge r_n \wedge \cdots \wedge r_{N+1})]$$

とあらわされ，次の再帰式が導かれる．

$$V^{N+1}(x, \lambda, \mu) = (\lambda \vee r_G(x)) - (\mu \wedge r_G(x)) \quad x \in X,\ \lambda \in \mathbf{R},\ \mu \in \mathbf{R}$$

$$V^n(x, \lambda, \mu) = \min_{u \in U} \sum_{y \in X} V^{n+1}(y\,,\ \lambda \vee r_n(x, u)\,,\ \mu \wedge r_n(x, u))p(y|x, u)$$
$$x \in X,\ \lambda \in \mathbf{R},\ \mu \in \mathbf{R},\quad n = 1, 2, \cdots, N$$

例2.4.4（分数型評価最大化問題）

Maximize $E^{\sigma}\left[\dfrac{r_1(x_1, u_1) + \cdots + r_N(x_N, u_N) + r_G(x_{N+1})}{R_1(x_1, u_1) + \cdots + R_N(x_N, u_N) + R_G(x_{N+1})}\right]$

subject to (i) $x_{n+1} \sim p(\cdot|x_n, u_n) \quad n = 1, 2, \ldots N$
(ii) $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_N\} \in \Sigma$

これは一般の複合型評価問題において

$$h(\xi, \eta) = \frac{\xi}{\eta},\quad \circ = +,\quad \diamond = +$$

とおいた場合にあたる．

分数型評価問題に対する部分問題群は

$$V^{N+1}(x_{N+1},\lambda,\mu) \;=\; \frac{\lambda + r_{N+1}}{\mu + R_{N+1}}$$

$$V^{n}(x_{n},\lambda,\mu) \;=\; \operatorname*{Max}_{\sigma\in\Sigma^n} E^{\sigma}\left[\frac{\lambda + r_n + \cdots + r_{N+1}}{\mu + R_n + \cdots + R_{N+1}}\right]$$

とあらわされ，次の再帰式が導かれる．

$$V^{N+1}(x,\lambda,\mu) = \frac{\lambda + r_G(x)}{\mu + R_G(x)} \quad x\in X,\ \lambda\in\mathbf{R},\ \mu\in\mathbf{R}$$

$$V^{n}(x,\lambda,\mu) = \operatorname*{Max}_{u\in U} \sum_{y\in X} V^{n+1}(y\,,\,\lambda + r_n(x,u)\,,\,\mu + R_n(x,u))p(y|x,u)$$

$$x\in X,\ \lambda\in\mathbf{R},\ \mu\in\mathbf{R},\quad n=1,2,\cdots,N$$

例2.4.5（分散型評価最小化問題）

ここでは，標本分散の最小化を考える．すなわち問題は以下のとおり

$$\text{minimize } E^{\sigma}\left[\frac{1}{N+1}\left\{\sum_{n=1}^{N+1}(r_n - \overline{r})^2\right\}\right]$$

$$\begin{aligned}\text{subject to }\ &\text{(i)}\ \ x_{n+1}\sim p(\,\cdot\,|x_n,u_n) \quad n=1,2,\ldots N\\ &\text{(ii)}\ \ \sigma=\{\sigma_1,\sigma_2,\ldots,\sigma_N\}\in\Sigma\end{aligned}$$

ただし

$$\overline{r} = \frac{1}{N+1}\sum_{n=1}^{N+1} r_n$$

とする．

ここで，目的関数を $(N+1)^2$ 倍すると

$$\begin{aligned}&(N+1)^2\times E^{\sigma}\left[\frac{1}{N+1}\left\{\sum_{n=1}^{N+1}(r_n-\overline{r})^2\right\}\right]\\ =\;&(N+1)^2\times E^{\sigma}\left[\frac{1}{N+1}\sum_{n=1}^{N+1} r_n^2 - \left(\frac{1}{N+1}\sum_{n=1}^{N+1} r_n\right)^2\right]\\ =\;&E^{\sigma}\left[\sum_{n=1}^{N+1}(N+1)r_n^2 - \left(\sum_{n=1}^{N+1} r_n\right)^2\right]\end{aligned}$$

したがって，一般の複合型評価問題において

$$h(\xi,\eta)=\xi-\eta^2,\quad \circ=+,\quad \diamond=+$$

とおき，$r_n=(N+1)r_n^2,\ R_n=r_n$ と置きなおした場合にあたる．

分散型評価問題は複合型評価問題とみなされたが，実際には2パラメータを必要とせず，1パラメータでよい．すなわち部分問題群

$$V^{N+1}(x_{N+1},\lambda) \;=\; (N+1)r_{N+1}^2 - (\lambda + r_{N+1})^2$$

$$V^{n}(x_n,\lambda) \;=\; \min_{\sigma\in\Sigma^n} E^{\sigma}\left[\sum_{m=n}^{N+1}(N+1)r_m^2 - \left(\lambda + \sum_{m=n}^{N+1} r_m\right)^2\right]$$

に対し，1パラメータの再帰式が導かれる．

$$V^{N+1}(x,\lambda) = (N+1)r_G^2(x) - (\lambda + r_G(x))^2 \quad x\in X,\ \lambda\in\mathbf{R}$$

$$V^{n}(x,\lambda) = \min_{u\in U}\sum_{y\in X}\Big[(N+1)r_n^2(x,u) + V^{n+1}(\,y\,,\lambda + r_n(x,u)\,)p(y|x,u)\Big]$$

$$x\in X,\quad \lambda\in\mathbf{R},\quad n=1,2,\cdots,N$$

例2.4.6（その他の評価）

成長確率最大化問題

$$\text{Maximize } P^{\pi}[r_1(x_1,u_1)\ \le\ \cdots\ \le\ r_N(x_N,u_N)\ \le\ r_G(x_{N+1})]$$

$$\begin{aligned}\text{subject to }\ &\text{(i)}\ \ x_{n+1}\sim p(\,\cdot\,|x_n,u_n) \quad n=1,2,\ldots N\\ &\text{(ii)}\ \ \pi=\{\pi_1,\pi_2,\ldots,\pi_N\}\in\Pi\end{aligned}$$

極値排除和最大化問題

$$\text{Maximize } E^{\sigma}[r_1+\cdots+r_{N+1} - (r_1\vee\cdots\vee r_{N+1}) - (r_1\wedge\cdots\wedge r_{N+1})]$$

$$\begin{aligned}\text{subject to }\ &\text{(i)}\ \ x_{n+1}\sim p(\,\cdot\,|x_n,u_n) \quad n=1,2,\ldots N\\ &\text{(ii)}\ \ \sigma=\{\sigma_1,\sigma_2,\ldots,\sigma_N\}\in\Sigma\end{aligned}$$