Examples of policies

(1) Always chose action 1 $\rightarrow$ let $D_n = 1 \quad \forall n$.

(2) If $X_n = a$ or $X_n = b$ then $D_n = 1$; otherwise $D_n = 2$

(3) If $X_n = a$ or $X_n = b$ then $D_n = 1$; If $X_n = c$ toss fair coin and if heads then $D_n = 1$ and if tails $D_n = 2$; if $X_n = d$ then $D_n = 2$

(4) Let $D_n = 1$ for $n = 0, 1$. For $n \geq 2$
if $X_{n-2} = a$ and $X_n > X_{n-1}$ then $D_n = 1$
if $X_{n-2} = b$ and $X_n > X_{n-1}$ and $D_{n-1} = 2$, then $D_n = 1$;
otherwise $D_n = 2$

stationary policy → Markov chain

Let a denote an action function

$$p^a(i,j) = p_{a(i)}(i,j) \quad \text{and} \quad f^a(i) = f_{a(i)}(i)$$

Example → $a = (1, 1, 2, 2)$

$$p^a = \begin{bmatrix} 0.1 & 0.3 & 0.6 & 0 \\ 0 & 0.2 & 0.5 & 0.3 \\ 0.8 & 0.2 & 0 & 0 \\ 0.9 & 0.1 & 0 & 0 \end{bmatrix} \qquad f^a = (100, 125, 350, 600)$$

Example $b = (1, 1, 1, 2)$

$$p^b = \begin{bmatrix} 0.1 & 0.3 & 0.6 & 0 \\ 0 & 0.2 & 0.5 & 0.3 \\ 0 & 0.1 & 0.2 & 0.7 \\ 0.9 & 0.1 & 0 & 0 \end{bmatrix} \qquad f^b = (100, 125, 150, 600)$$

Saturday, March 21, 2020 11:14 AM

$\pi P = \pi$

$\nu P = \nu$

invariant function unique → fixed point

Find $x \ni x = f(x)$

sometimes step 1. Let x_0 be arbitrary

step 2. Let $x_{n+1} = f(x_n)$

$\downarrow$

$x = f(x)$

Find $x \ni x = e^{-x}$

let $x_0 = 0$

$x_1 = e^{-0} = 1$

$x_2 = e^{-1} = 0.3679$

$x_3 = e^{-0.3679} = 0.6922$

$x_4 = e^{-0.6922} = 0.5005$

$x_5 = e^{-0.5005} = 0.6062$

$\vdots$

$x^* = 0.5671\ldots$

$T(f) = g$ where T is a transformation mapping functions into functions.

h is a fixed point of T if $T(h) = h$

Sometimes let $h_0 \equiv 0$, $h_{n+1} = T(h_n)$

$T(V_{old}) = V_{new}$

$$V_{new}(i) = \min_{k \in A} \left\{ f_k(i) + \alpha \sum_{j \in E} P_k(i,j) V_{old}(j) \right\}$$

$v^{\alpha} = (4287.4,\ 4381.6,\ 4440.9,\ 4612.9)^T$ for $\alpha = 0.95$

$$v^{\alpha}(a) = \min\left\{ 100 + 0.95\,(0.1,\ 0.3,\ 0.6,\ 0)\begin{pmatrix} 4287.4 \\ 4381.6 \\ 4440.9 \\ 4612.9 \end{pmatrix};\ \rightarrow k=1 \right.$$

$$\left. 300 + 0.95\,(0.6,\ 0.3,\ 0.1,\ 0)\begin{pmatrix} 4287.4 \\ 4381.6 \\ 4440.9 \\ 4612.9 \end{pmatrix} \right\} \rightarrow k=2$$

$$= \min\{4287.4,\ 4414.5\} = 4287.4$$

Action function

$a^*(a) = 1$

$$v^{\alpha}(b) = \min\left\{ 125 + 0.95\,(0,\ 0.2,\ 0.5,\ 0.3)\begin{pmatrix} 4287.4 \\ 4381.6 \\ 4440.9 \\ 4612.9 \end{pmatrix};\right.$$

$$\left. 325 + 0.95\,(0.75,\ 0.1,\ 0.1,\ 0.05)\begin{pmatrix} 4287.4 \\ 4381.6 \\ 4440.9 \\ 4612.9 \end{pmatrix} \right\}$$

$a^*(b) = 1$ $\quad = \min\{4381.6,\ 4437.0\} = 4381.6$

$v^{\alpha}(c) = \cdots = \min\{4477.6,\ 4440.9\} = 4440.9$

$a^*(c) = 2$

$v^{\alpha}(d) = \cdots = \min\{4612.9,\ 4682.0\} = 4612.9$

$a^*(d) = 1 \Rightarrow a^* = (1, 1, 2, 1)$

$v_0 = (0, 0, 0, 0)$

$V_1 = (100, 125, 150, 500)$

$V_2 = (230.62, 362.5, 449.75, 635.38)$

$V_3 = \ldots$

$\alpha = 0.95$

Iteration $n=0$

Step 1. $a_0 = (1, 1, 1, 1)$

Step 2. $f = \begin{pmatrix} 100 \\ 125 \\ 150 \\ 500 \end{pmatrix}$, $P = \begin{bmatrix} 0.1 & 0.3 & 0.6 & 0 \\ 0 & 0.2 & 0.5 & 0.3 \\ 0 & 0.1 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0 & 0.1 \end{bmatrix}$

Step 3 $I - \alpha P = \begin{bmatrix} 0.905 & -0.285 & -0.57 & 0 \\ 0 & 0.81 & -0.475 & -0.285 \\ 0 & -0.095 & 0.81 & -0.665 \\ -0.76 & -0.095 & 0 & 0.905 \end{bmatrix}$

$(I - \alpha P)^{-1} f = \begin{pmatrix} 4502 \\ 4591 \\ 4676 \\ 4815 \end{pmatrix}$

Step 4.

$$a_1(a) = \operatorname{argmin}\left\{ 100 + 0.95 (0.1, 0.3, 0.6, 0) \begin{pmatrix} 4502 \\ 4591 \\ 4676 \\ 4815 \end{pmatrix}; \right.$$
$$\left. 300 + 0.95 (0.6, 0.3, 0.1, 0) \begin{pmatrix} 4502 \\ 4591 \\ 4676 \\ 4815 \end{pmatrix} \right\} = \operatorname{argmin}\{4501.4, 4618.8\}$$
$a_1(a) = 1$

$$a_1(b) = \operatorname{argmin}\left\{ 125 + 0.95 (0, 0.2, 0.5, 0.3) \begin{pmatrix} 4502 \\ 4591 \\ 4676 \\ 4815 \end{pmatrix}; \right.$$
$$\left. 325 + 0.95 (0.75, 0.1, 0.1, 0.05) \begin{pmatrix} 4502 \\ 4591 \\ 4676 \\ 4815 \end{pmatrix} \right\} = \operatorname{argmin}\{4590.7, 4641.8\}$$
$a_1(b) = 1$

$a_1(c) = \dots = \operatorname{argmin}\{4676.6, 4643.8\}$ $a_1(c) = 2$

$a_1(d) = \dots = \arg\min \{4815.1, 4885.4\}$ $a_1(d) = 1$

$a_0 = (1, 1, 1, 1)^T$ and $a_1 = (1, 1, 2, 1)^T$ repeat steps 2, 3, 4