

## Auxiliar 8: Markov con Decisiones

Martes 26 de Octubre de 2010

### Problema 1

a) La cadena queda se muestra en la figura.

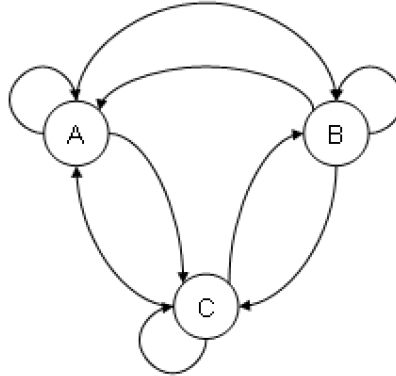


Figura 1: Cadena problema 1

Las políticas de decisión pueden ser especificadas como combinaciones de las políticas puras de publicidad y no publicidad. Sólo basta con especificar las matrices de transición para estas últimas (ver el enunciado).

Respecto a los beneficios asociados a las políticas estos serán:

$$r^{Sin} = \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix} \quad y \quad r^{Pub} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$$

b) Esto no es más que resolver una programación dinámica donde el estado son las ventas y las decisiones son hacer o no publicidad. Esto queda de la siguiente forma (ojo, que el número de estados es 3 por lo que podemos especificar la función de beneficios para cada uno):

**Etapas 0** (contando desde el final hacia atrás):

$$V(0) = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$$

**Etapas k:**

$$V_A(k) = \max \left\{ 3 + (0,5 \quad 0,3 \quad 0,2) \cdot \begin{pmatrix} V_A(k-1) \\ V_M(k-1) \\ V_B(k-1) \end{pmatrix}, 5 + (0,2 \quad 0,5 \quad 0,3) \cdot \begin{pmatrix} V_A(k-1) \\ V_M(k-1) \\ V_B(k-1) \end{pmatrix} \right\}$$

$$V_M(k) = \max \left\{ 1 + (0,4 \quad 0,4 \quad 0,2) \cdot \begin{pmatrix} V_A(k-1) \\ V_M(k-1) \\ V_B(k-1) \end{pmatrix}, 3 + (0,1 \quad 0,4 \quad 0,5) \cdot \begin{pmatrix} V_A(k-1) \\ V_M(k-1) \\ V_B(k-1) \end{pmatrix} \right\}$$

$$V_B(k) = \max \left\{ -1 + (0,4 \quad 0,6 \quad 0,0) \cdot \begin{pmatrix} V_A(k-1) \\ V_M(k-1) \\ V_B(k-1) \end{pmatrix}, 1 + (0,0 \quad 0,3 \quad 0,7) \cdot \begin{pmatrix} V_A(k-1) \\ V_M(k-1) \\ V_B(k-1) \end{pmatrix} \right\}$$

c) Utilizaremos el algoritmo de Howard:

$$\bar{S} = \begin{pmatrix} A \rightarrow Sin \\ M \rightarrow Sin \\ B \rightarrow Pub \end{pmatrix}$$

La matriz asociada a esta política es la siguiente:

$$\bar{M} = \begin{pmatrix} 0,2 & 0,5 & 0,3 \\ 0,1 & 0,4 & 0,5 \\ 0,4 & 0,6 & 0,0 \end{pmatrix} \quad y \quad \bar{r} = \begin{pmatrix} 5 \\ 3 \\ -1 \end{pmatrix}$$

Ahora calculamos el vector  $\bar{W}$ :

$$\bar{W} + g \cdot \vec{e} = \bar{r} + \bar{M} \cdot \bar{W}$$

Desde  $\Pi = \Pi \cdot \bar{M}$  y  $\sum_i \Pi_i = 1$  tenemos que:

$$\Pi = \begin{pmatrix} 0,22 \\ 0,48 \\ 0,30 \end{pmatrix}$$

Entonces:

$$g = \sum_i \Pi_i \cdot \bar{r}_i = 2,2$$

Por lo tanto:

$$(I - \bar{M})\bar{W} = \bar{r} - g \cdot e$$

$$\begin{pmatrix} 0,8 & -0,5 & -0,3 \\ -0,1 & 0,6 & -0,5 \\ -0,4 & -0,6 & 1,0 \end{pmatrix} \cdot \begin{pmatrix} W_A \\ W_M \\ W_B \end{pmatrix} = \begin{pmatrix} 2,79 \\ 0,79 \\ -3,2 \end{pmatrix} \Rightarrow \begin{matrix} W_A = 5,2 \\ W_M = 2,4 \\ W_B = 0,0 \end{matrix}$$

Ahora debemos construir la próxima política estacionaria.

$$S(A) = \begin{cases} 5 + (0,2 \quad 0,5 \quad 0,3) \cdot \begin{pmatrix} 5,2 \\ 2,4 \\ 0 \end{pmatrix} = 7,2 \\ 3 + (0,5 \quad 0,3 \quad 0,2) \cdot \begin{pmatrix} 5,2 \\ 2,4 \\ 0 \end{pmatrix} = 6,34 \end{cases} \Rightarrow Sin$$

$$S(M) = \begin{cases} 3 + (0,1 \quad 0,4 \quad 0,5) \cdot \begin{pmatrix} 5,2 \\ 2,4 \\ 0 \end{pmatrix} = 4,5 \\ 1 + (0,4 \quad 0,4 \quad 0,2) \cdot \begin{pmatrix} 5,2 \\ 2,4 \\ 0 \end{pmatrix} = 4,06 \end{cases} \Rightarrow Sin$$

$$S(B) = \begin{cases} 1 + (0,0 \quad 0,3 \quad 0,7) \cdot \begin{pmatrix} 5,2 \\ 2,4 \\ 0 \end{pmatrix} = 1,7 \\ -1 + (0,4 \quad 0,6 \quad 0,0) \cdot \begin{pmatrix} 5,2 \\ 2,4 \\ 0 \end{pmatrix} = 2,5 \end{cases} \Rightarrow Pub$$

Pero hemos reconstruido la política  $\bar{S} \Rightarrow \bar{S}$  es la política óptima.

## Problema 2

- a) ■ Estados: niveles posibles de stock ( $X_t$ )

$$E = \{0, 1, \dots, C\}$$

- Decisiones: número de unidades ordenadas al principio del día ( $Y_t - X_t$ )

$$A_i = \{0, 1, \dots, C - i\} \quad \text{con } i \in E$$

- Probabilidades de transición:

$$p_{ij}(a) = \mathcal{P}[X_{t+1} = j \mid X_t = i, a]$$

- caso  $j > i + a$ :  $p_{ij}(a) = 0$
- caso  $j > 0$  y  $j > i + a$ :  $p_{ij}(a) = 0$
- caso  $0 \leq j \leq i + a$ :

$$\begin{aligned} p_{ij}(a) &= \mathcal{P}[Y_t - D_t = j \mid Y_t = i + a, a] \\ &= \mathcal{P}[D_t = i + a - j \mid Y_t = i + a, a] \\ &= q_{i+a-j, i+a} \end{aligned}$$

- caso  $j = 0$ :

$$\begin{aligned} p_{ij}(a) &= \mathcal{P}[Y_t - D_t \leq 0 \mid Y_t = i + a, a] \\ &= \mathcal{P}[N \geq D_t \geq i + a \mid Y_t = i + a, a] \\ &= \sum_{k=0}^{N-i-a} q_{i+a+k, i+a} \end{aligned}$$

En resumen:

$$p_{ij}(a) = \begin{cases} q_{i+a-j, i+a} & \text{si } 0 < j \leq i + a \\ \sum_{k=0}^{N-i-a} q_{i+a+k, i+a} & \text{si } j = 0 \\ 0 & \text{si no} \end{cases}$$

- Función de Beneficios:

$$\begin{aligned} \hat{r}_i(a) &= \mathbb{E}[v \min\{D_t, Y_t\} \mid X_t = i \wedge a = Y_t - X_t] - \mathbb{E}[h(Y_t - D_t) \mid X_t = i \wedge a = Y_t - X_t] - ca \\ &= v \left[ \sum_{d=0}^{i+a} dq_{d, i+a} + (i+a) \cdot \sum_{d=i+a+1}^N q_{d, i+a} \right] - h \sum_{d=0}^{i+a} (i+a-d)q_{d, i+a} - ca \end{aligned}$$

- b) Las matrices de transición para cada decisión son:

$$P(0) = \begin{pmatrix} 1 & 0 & 0 \\ 0,9 & 0,1 & 0 \\ 0,6 & 0,4 & 0 \end{pmatrix} \quad P(1) = \begin{pmatrix} 0,9 & 0,1 & 0 \\ 0,6 & 0,4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad P(2) = \begin{pmatrix} 0,6 & 0,4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Donde por ejemplo

$$\begin{aligned}
p_{10}(0) &= \text{Probabilidad de pasar de un stock de 1 a uno de 0 dado que no ordena nada al principio del día} \\
&= q_{11} + q_{21} \\
&= 0,4 + 0,5 \\
&= 0,9
\end{aligned}$$

Utilizando la fórmula de los beneficios encontrada en la parte a) se calcula:

$$\hat{r}(0) = \begin{pmatrix} 0 \\ 5,1 \\ 8,4 \end{pmatrix} \quad \hat{r}(1) = \begin{pmatrix} 3,1 \\ 6,4 \\ 0 \end{pmatrix} \quad \hat{r}(2) = \begin{pmatrix} 4,4 \\ 0 \\ 0 \end{pmatrix}$$

La política estacionaria con la que iniciaremos el algoritmo es la que siempre elige tener stock de 2, es decir:

$$s = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

Luego planteamos el sistema:

$$W^s + g^s e = \hat{r}^s + P^s W^s$$

$$\text{con } P^s = \begin{pmatrix} 0,6 & 0,4 & 0 \\ 0,6 & 0,4 & 0 \\ 0,6 & 0,4 & 0 \end{pmatrix} \text{ y } \hat{r}^s = \begin{pmatrix} 4,4 \\ 6,4 \\ 8,4 \end{pmatrix}$$

El sistema es entonces:

$$\begin{aligned}
w_0 &= 4,4 - g + 0,6w_0 + 0,4w_1 \\
w_1 &= 6,4 - g + 0,6w_0 + 0,4w_1 \\
w_2 &= 8,4 - g + 0,6w_0 + 0,4w_1
\end{aligned}$$

Se impone  $w_2 = 0$  y se llega a  $w = \begin{pmatrix} -4 \\ -2 \\ 0 \end{pmatrix}$  y  $g = 5,2$ . Luego recalculamos la política:

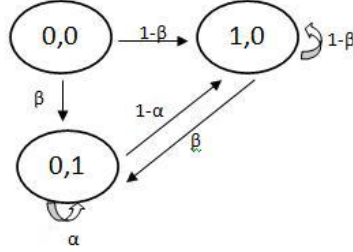
$$\bar{s}(i) = \arg \max_{a \in A_i} \left\{ \hat{r}_i(a) + \sum_j p_{ij}(a) W_j^s \right\}$$

Estado	Decisión	Valor
0	0	$0 + 1 \cdot (-4) = -4$
0	1	$3,1 + 0,9 \cdot (-4) + 0,1 \cdot (-2) = -0,7$
0	2	$4,4 + 0,6 \cdot (-4) + 0,4 \cdot (-2) = \boxed{1,2}$
1	0	$5,1 + 0,9 \cdot (-4) + 0,1 \cdot (-2) = 1,3$
1	1	$6,4 + 0,6 \cdot (-4) + 0,4 \cdot (-2) = \boxed{3,2}$

La nueva política es  $\bar{s} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ . Como  $s = \bar{s}$  se termina el algoritmo.

### Problema 3

- (a) La capacidad óptima de la bodega es uno, pues dado que la demanda es binaria y existe costo de almacenamiento, nunca se va a querer producir cuando exista un producto en bodega. *OBS: También pueden poner el caso que es cero cuando el costo de producción es mayor al precio de venta*
- (b) Cada estado es un par ordenado (i,j) que representa el sistema AL COMIENZO de la semana. Se tiene que i señala si existe o no el producto en bodega y j señala la realización de la demanda la semana anterior.



De aquí en adelante  $A=(0,0)$ ;  $B=(0,1)$  y  $C=(1,0)$

- (c) Por la parte (a) ya se vio que  $s_t^*(C) = \text{No producir} \forall t$  Por otra parte, por enunciado se supone que:

$$\begin{aligned} V_0^*(A) &= 0 \\ V_0^*(B) &= 0 \\ V_0^*(C) &= 100 \end{aligned}$$

Luego,

$$\begin{aligned} V_1^*(A) &= \max\left\{ \begin{pmatrix} 1-\beta & \beta & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 100 \end{pmatrix}, \begin{pmatrix} 0 & \beta & 1-\beta \end{pmatrix} \begin{pmatrix} P-c \\ P-c \\ -c+100 \end{pmatrix} \right\} \\ &= \max\{0, \beta(P-c) + (1-\beta)(-c+100)\} \\ &= \max\{0, (P-c)\beta\} \\ &= \max\{0, 100 * 0,9\} \\ &= \max\{0, 90\} \end{aligned}$$

Por lo tanto,  $s_1^*(A) = \{\text{Producir}\}$

En tanto para B se tiene,

$$\begin{aligned} V_1^*(B) &= \max\left\{ \begin{pmatrix} 1-\alpha & \alpha & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 100 \end{pmatrix}, \begin{pmatrix} 0 & \alpha & 1-\alpha \end{pmatrix} \begin{pmatrix} P-c \\ P-c \\ -c+h \end{pmatrix} \right\} \\ &= \max\{0, \alpha(P-c) + (1-\alpha)(-c+100)\} \\ &= \max\{0, (P-c) * \alpha\} \\ &= \max\{0, 100 * 0,2\} \\ &= \max\{0, 20\} \end{aligned}$$

Por lo tanto,  $s_1^*(B) = \{\text{Producir}\}$

- (d) **Paso 1**

Se parte con la política  $\bar{s}=(P,P,NP)$ . Por lo tanto, se tiene que:

$$P^{\bar{s}} = \begin{pmatrix} 0 & 0,9 & 0,1 \\ 0 & 0,2 & 0,8 \\ 0 & 0,9 & 0,1 \end{pmatrix}, R^{\bar{s}} = \begin{pmatrix} 0 & 0 & -c-h \\ 0 & P-c & -c-h \\ 0 & P & -h \end{pmatrix}$$

luego,

$$R^{\bar{s}} = \begin{pmatrix} 0 & 0 & -150 \\ 0 & 100 & -150 \\ 0 & 200 & -50 \end{pmatrix}$$

$$\hat{r} = \begin{pmatrix} -150 \cdot 0,1 \\ 100 \cdot 0,2 - 150 \cdot 0,8 \\ 200 \cdot 0,9 - 50 \cdot 0,1 \end{pmatrix} = \begin{pmatrix} -15 \\ -100 \\ 175 \end{pmatrix}$$

### Paso 2

Obtenemos  $g^{\bar{s}}$

$$\begin{aligned} \pi_B &= 0,2\pi_B + 0,9\pi_C \\ \pi_C &= 0,8\pi_B + 0,1\pi_C \\ \pi_B + \pi_C &= 1 \end{aligned}$$

Además  $\pi_A = 0$ , pues es transiente.

$$\begin{aligned} \pi_A &= 0 \\ \pi_B &= 0,529 \\ \pi_C &= 0,471 \end{aligned}$$

Por lo tanto,  $g^{\bar{s}} = 29,525$

Obtenemos  $w^{\bar{s}}$

$$\begin{aligned} w + ge &= r + Pw \\ (I - P)w &= r - ge \\ \begin{pmatrix} 1 & -0,9 & -0,1 \\ 0 & 0,8 & -0,8 \\ 0 & -0,9 & 0,9 \end{pmatrix} \begin{pmatrix} w_A \\ w_B \\ 0 \end{pmatrix} &= \begin{pmatrix} -44,525 \\ -129,525 \\ 145,475 \end{pmatrix} \\ \begin{pmatrix} w_A \\ w_B \\ w_C \end{pmatrix} &= \begin{pmatrix} -190,241 \\ -161,906 \\ 0 \end{pmatrix} \end{aligned}$$

### Paso 3

Obtener la política estacionaria s

$$\begin{aligned} s_A &= \arg \max_{a \in \{P, NP\}} \{r_A(NP) + \sum_j p_{A,j}(NP)w_j^{\bar{s}}, r_A(P) + \sum_j p_{A,j}(P)w_j^{\bar{s}}\} \\ &= \arg \max_{a \in \{P, NP\}} \{0 + 0,1 \cdot (-190,241) + 0,9 \cdot (-161,906) + 0, -15 + 0,9 \cdot (-190,241) + 0,1 \cdot 0\} \\ &= \arg \max_{a \in \{P, NP\}} \{-141,982; -186,217\} \\ &= NP \\ s_B &= \arg \max_{a \in \{P, NP\}} \{r_B(NP) + \sum_j p_{B,j}(NP)w_j^{\bar{s}}, r_B(P) + \sum_j p_{B,j}(P)w_j^{\bar{s}}\} \\ &= \arg \max_{a \in \{P, NP\}} \{20 + 0,8 \cdot (-190,241) + 0,2 \cdot (-161,906) + 0,100 + 0,2 \cdot (-161,241)\} \\ &= \arg \max_{a \in \{P, NP\}} \{-164,574; 67,75\} \\ &= P \end{aligned}$$

Por lo tanto como la nueva política NO es igual a la anterior, la política NO es óptima.