



## Problema Markov con decisiones

Viernes 9 de Julio de 2010

### Problema 1

En una granja del sur de Chile se quiere encontrar la política óptima de reemplazo de vacas lecheras. Al final de cada período semestral se observa la productividad de una vaca, la que puede ser alta, media, o baja. Los ingresos naturalmente dependen de la productividad y son de 0, 100, 200 mil pesos si la productividad es baja, media o alta respectivamente. Al final de cada período se puede decidir mantener o reemplazar la vaca. Si la vaca se mantiene, su productividad en el siguiente período se distribuye de acuerdo a la siguiente tabla:

	<i>B</i>	<i>M</i>	<i>A</i>
<i>B</i>	0,6	0,3	0,1
<i>M</i>	0,2	0,6	0,2
<i>A</i>	0,1	0,3	0,6

Por otra parte, reemplazar la vaca tiene un costo asociado de 100 mil pesos, y la productividad de una vaca nueva se distribuye equiprobablemente, es decir, su productividad es alta, media, o baja con probabilidad  $\frac{1}{3}$ .

- a) Modele la situación como un proceso de decisión markoviano donde se quiere maximizar las ganancias medias en cada periodo, suponiendo un horizonte finito de tiempo.

Estados = productividad de la vaca = {Baja, Media, Alta}

Acciones = {Mantener, Reemplazar}

Matrices de transición según las acciones:

$$P^M = \begin{pmatrix} 0,6 & 0,3 & 0,1 \\ 0,2 & 0,6 & 0,2 \\ 0,1 & 0,3 & 0,6 \end{pmatrix}, P^R = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Los beneficios según las acciones son:

$$r^M = \begin{pmatrix} 0 \\ 100 \\ 200 \end{pmatrix}, r^R = \begin{pmatrix} -100 \\ 0 \\ 100 \end{pmatrix}$$

Luego en cada periodo, dado que faltan  $k$  periodos en adelante, y se está en el estado  $i$ , se resuelve el problema:

$$V_k(i) = \max_{a \in M, R} \left\{ \sum_{j \in B, M, A} r_i(a) + [r_{i,j}(a) + V_j(k+1)] p_{i,j}(a) \right\}$$

$$V_k(i) = \max\{V_k^M(i), V_k^R(i)\}$$

Donde:

$$V_k^M = \begin{pmatrix} 0 \\ 100 \\ 200 \end{pmatrix} + \begin{pmatrix} 0,6 & 0,3 & 0,1 \\ 0,2 & 0,6 & 0,2 \\ 0,1 & 0,3 & 0,6 \end{pmatrix} \begin{pmatrix} V_{k-1}(B) \\ V_{k-1}(M) \\ V_{k-1}(A) \end{pmatrix}$$

$$V_k^R = \begin{pmatrix} -100 \\ 0 \\ 100 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} V_{k-1}(B) \\ V_{k-1}(M) \\ V_{k-1}(A) \end{pmatrix}$$

- b) Usando el algoritmo de Howard encuentre la política óptima de reemplazo de vacas, y calcule las ganancias medias por período en el largo plazo.

### Paso 1

Partamos con la política inicial  $\bar{s} = (M, R, R)$ . Luego se tiene que:

$$P^{\bar{s}} = \begin{pmatrix} 0,6 & 0,3 & 0,1 \\ 0,3 & 0,3 & 0,3 \\ 0,3 & 0,3 & 0,3 \end{pmatrix}, r^{\bar{s}} = \begin{pmatrix} 0 \\ 0 \\ 100 \end{pmatrix}$$

### Paso 2

Obtenemos  $g^{\bar{s}}$ , (para luego obtener  $w^{\bar{s}}$ ):

$$\begin{aligned} \pi^T &= \pi^T P \\ \pi_B + \pi_M + \pi_A &= 1 \end{aligned}$$

$$\begin{aligned} 0,4\pi_B - 0,3\pi_M - 0,3\pi_A &= 0 \\ -0,3\pi_B + 0,6\pi_M - 0,3\pi_A &= 0 \\ -0,1\pi_B - 0,3\pi_M + 0,6\pi_A &= 0 \\ \pi_B + \pi_M + \pi_A &= 1 \end{aligned}$$

$$\begin{aligned} \pi^T &= (0,45, 0,318, 0,227) \\ g^{\bar{s}} &= 22,72 \end{aligned}$$

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

$$\begin{aligned} w + ge &= r + Pw \\ (I - P)w &= r - ge \\ \begin{pmatrix} 0,4 & -0,3 & -0,1 \\ -0,3 & 0,6 & -0,3 \\ -0,3 & -0,3 & 0,6 \end{pmatrix} \begin{pmatrix} w_B \\ w_M \\ 0 \end{pmatrix} &= \begin{pmatrix} -22,72 \\ -22,72 \\ 77,27 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} 0,4w_B - 0,3w_M &= -22,72 \\ -0,3w_B + 0,6w_M &= -22,72 \end{aligned}$$

$$w^{\bar{s}} = \begin{pmatrix} w_B \\ w_M \\ w_A \end{pmatrix} = \begin{pmatrix} -131,81 \\ -100 \\ 0 \end{pmatrix}$$

### Paso 3

Obtener la política estacionaria  $s$  de la forma que cumpla:

$$s(E_i) = \arg \max_{a \in M, R} \left\{ r_i(a) + \sum_j p_{i,j}(a) w_j^{\bar{s}} \right\}$$

$E_i = B$ :

$$\begin{aligned} s_B &= \arg \max_{a \in M, R} \left\{ 0 + \begin{pmatrix} 0,6 & 0,3 & 0,1 \end{pmatrix} \begin{pmatrix} -131,81 \\ -100 \\ 0 \end{pmatrix}, -100 + \begin{pmatrix} 0,3 & 0,3 & 0,3 \end{pmatrix} \begin{pmatrix} -131,81 \\ -100 \\ 0 \end{pmatrix} \right\} \\ &= \arg \max_{a \in M, R} \{-109,09, -177,27\} \\ &= M \end{aligned}$$

$$E_i = M:$$

$$\begin{aligned} s_M &= \arg \max_{a \in M, R} \left\{ 100 + \begin{pmatrix} 0,2 & 0,6 & 0,2 \end{pmatrix} \begin{pmatrix} -131.\overline{81} \\ -100 \\ 0 \end{pmatrix}, 0 + \begin{pmatrix} 0.\overline{3} & 0.\overline{3} & 0.\overline{3} \end{pmatrix} \begin{pmatrix} -131.\overline{81} \\ -100 \\ 0 \end{pmatrix} \right\} \\ &= \arg \max_{a \in M, R} \{13.\overline{63}, -77.\overline{27}\} \\ &= M \end{aligned}$$

$$E_i = A:$$

$$\begin{aligned} s_A &= \arg \max_{a \in M, R} \left\{ 200 + \begin{pmatrix} 0,1 & 0,3 & 0,6 \end{pmatrix} \begin{pmatrix} -131.\overline{81} \\ -100 \\ 0 \end{pmatrix}, 100 + \begin{pmatrix} 0.\overline{3} & 0.\overline{3} & 0.\overline{3} \end{pmatrix} \begin{pmatrix} -131.\overline{81} \\ -100 \\ 0 \end{pmatrix} \right\} \\ &= \arg \max_{a \in M, R} \{156.\overline{81}, 22.\overline{72}\} \\ &= M \end{aligned}$$

Luego  $s = (M, M, M)$ .

#### Paso 4

Como  $s \neq \bar{s}$ , entonces iteramos de nuevo desde el paso 2.

#### Segunda iteración:

##### Paso 2

Tomamos como politica estacionaria  $\bar{s} = (M, M, M)$ .

Obtenemos  $g^{\bar{s}}$ , (para luego obtener  $w^{\bar{s}}$ ):

$$\begin{aligned} \pi^T &= \pi^T P \\ \pi_B + \pi_M + \pi_A &= 1 \end{aligned}$$

$$\begin{aligned} 0,4\pi_B - 0,2\pi_M - 0,1\pi_A &= 0 \\ -0,3\pi_B + 0,4\pi_M - 0,3\pi_A &= 0 \\ -0,1\pi_B - 0,2\pi_M + 0,4\pi_A &= 0 \\ \pi_B + \pi_M + \pi_A &= 1 \end{aligned}$$

$$\pi^T = (0.\overline{285714}, 0.\overline{428571}, 0.\overline{285714})$$

$$g^{\bar{s}} = 100$$

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

$$w + ge = r + Pw$$

$$(I - P)w = r - ge$$

$$\begin{pmatrix} 0,4 & -0,3 & -0,1 \\ -0,2 & 0,4 & -0,2 \\ -0,1 & -0,3 & 0,4 \end{pmatrix} \begin{pmatrix} w_B \\ w_M \\ 0 \end{pmatrix} = \begin{pmatrix} -100 \\ 0 \\ 100 \end{pmatrix}$$

$$\begin{aligned} 0,4w_B - 0,3w_M &= -100 \\ -0.\overline{3}w_B + 0.\overline{6}w_M &= 0 \end{aligned}$$

$$w^{\bar{s}} = \begin{pmatrix} w_B \\ w_M \\ w_A \end{pmatrix} = \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix}$$

**Paso 3**

Obtener la política estacionaria  $s$  de la forma que cumpla:

$$s(E_i) = \arg \max_{a \in M, R} \left\{ r_i(a) + \sum_j p_{i,j}(a) w_j^{\bar{s}} \right\}$$

$E_i = B$ :

$$\begin{aligned} s_B &= \arg \max_{a \in M, R} \left\{ 0 + \begin{pmatrix} 0,6 & 0,3 & 0,1 \end{pmatrix} \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix}, -100 + \begin{pmatrix} 0,3 & 0,3 & 0,3 \end{pmatrix} \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix} \right\} \\ &= \arg \max_{a \in M, R} \{-300, -300\} \\ &= M \end{aligned}$$

$E_i = M$ :

$$\begin{aligned} s_M &= \arg \max_{a \in M, R} \left\{ 100 + \begin{pmatrix} 0,2 & 0,6 & 0,2 \end{pmatrix} \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix}, 0 + \begin{pmatrix} 0,3 & 0,3 & 0,3 \end{pmatrix} \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix} \right\} \\ &= \arg \max_{a \in M, R} \{-100, -200\} \\ &= M \end{aligned}$$

$E_i = A$ :

$$\begin{aligned} s_A &= \arg \max_{a \in M, R} \left\{ 200 + \begin{pmatrix} 0,1 & 0,3 & 0,6 \end{pmatrix} \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix}, 100 + \begin{pmatrix} 0,3 & 0,3 & 0,3 \end{pmatrix} \begin{pmatrix} -400 \\ -200 \\ 0 \end{pmatrix} \right\} \\ &= \arg \max_{a \in M, R} \{100, -100\} \\ &= M \end{aligned}$$

Luego  $s = (M, M, M)$ .

**Paso 4**

Como  $s = \bar{s}$ , entonces terminamos, y la política óptima es  $s = (M, M, M)$ .