Processos de Decisão Markovianos

<u>Processo de Decisão Markoviano (PDM)</u> $\Rightarrow$ Processo Estocástico no qual o estado do processo no futuro depende apenas do estado do processo e da decisão escolhida no presente. Em estado e tempo discretos fica:

$$P\{X(t+1)=x_{t+1} \mid X(t)=x_t, d_{x_t}(R)=k_t, X(t-1)=x_{t-1}, d_{x_{t-1}}(R)=k_{t-1}, \ldots, X(0)=x_0, d_0(R)=k_0\}=$$

$$P\{X(t+1)=x_{t+1} \mid X(t)=x_t, d_{x_t}(R)=k_t\}$$

<u>PDM é descrito por 4 tipos de informações</u>:

1. Espaço de Estados $\Rightarrow$ $E=\{0,1,\ldots,M\}$
2. Conjunto de Decisões $\Rightarrow$ para cada estado i, existe um conjunto de decisões possíveis $d_i(R)=\{1.2,\ldots K\}$, sendo tomada apenas uma delas segundo uma política R. Assim $\{\Delta_t, t=0,1,\ldots\}$é a seqüência de decisões tomadas. R é $\{d_0(R), d_1(R), \ldots, d_M(R)\}$.
3. Probabilidades de Transição $\Rightarrow p_{ij}(k)=P\{X(t+1)=j \mid X(t)=i, d_i(R)=k\}$
4. Custos Esperados $\Rightarrow$ C_{ik} é o custo esperado de se tomar a decisão k com o processo no estado i.

PDM $\Rightarrow$ seqüência de estados X(0), X(1),.... e decisões tomadas Δ_0, Δ_1,....

<u>Objetivo Principal</u> $\Rightarrow$ Determinar a política R que minimize os custos a longo período (horizonte infinito). Obs: Problemas de Programação Dinâmica Não Determinísticos com Horizonte Infinito = PDM.

Exemplo Protótipo

Uma máquina engarrafadora de água em perfeitas condições pode no dia seguinte apresentar algum defeito com probabilidade 0.09 ou passar a uma situação de avaria total com probabilidade 0,01. Trabalhando com defeito, a máquina pode manter-se neste estado no dia seguinte com probabilidade 0,55 ou passar ao estado de avaria total com probabilidade 0,45.

Espaço de Estados $\Rightarrow E = \{0,1,2\} \Rightarrow \{\text{perfeito}, \text{defeito}, \text{avaria\_total}\}$

Matriz de Transição $\Rightarrow$

$$P = \begin{array}{c c} \text{Estado} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 0.9 & 0.09 & 0.01 \\ 0 & 0.55 & 0.45 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

Considerando as seguintes ações possíveis para tomada de decisão:

Decisão	Ação
1	Não fazer nada
2	Reparar a máquina
3	Substituir a máquina

Algumas políticas possíveis com as decisões dadas são:

Política	$d_0(R)$	$d_1(R)$	$d_2(R)$
R_1	1	1	3
R_2	1	3	3
R_3	1	2	3
R_4	1	2	2

A substituição faz com que o processo retorne ao estado 0, mas a reparação se for feita no estado 1, em 80 % dos casos a máquina fica em perfeitas condições e no restante fica na mesma situação com defeito. Se a reparação for feita no estado 2, em 30 % dos casos a máquina fica em perfeitas condições e no restante fica na mesma situação com defeito. As matrizes de transição para cada política são:

$$P(R_1) = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.9 & 0.09 & 0.01 \\ 0 & 0.55 & 0.45 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

$$P(R_2) = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.9 & 0.09 & 0.01 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

$$P(R_3) = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.9 & 0.09 & 0.01 \\ 0.8 & 0.2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

$$P(R_4) = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.9 & 0.09 & 0.01 \\ 0.8 & 0.2 & 0 \\ 0.3 & 0.7 & 0 \end{bmatrix} \end{array}$$

Considerando que a reparação da máquina custa \$30,00 e perda de um dia de produção, a substituição custa \$60,00 e perda de um dia de produção, funcionando com defeito custa \$10,00 e o custo de um dia de produção perdida é \$20,00 os custos para cada estado são:

Estado	C_{ik} Decisão		
	1	2	3
0	0	50	80
1	10	50	80
2	inf	50	80

Pode-se agora calcular o custo de cada política através de:

$$E(C) = \sum_{i=0}^{M} C_{ik} \pi_i$$

Política	(π_0, π_1, π_2)	E(C)	
R_1	(0.7692,0.1539,0.0769)	7.6923	
R_2	(0.9091,0.0818,0.0091)	7.2727	
R_3	(0.8909,0.1002,0.0089)	5.7238	Mínimo
R_4	(0.8840,0.1072,0.0088)	5.8011	

<u>Determinação de Políticas Ótimas via Programação Linear</u> $\Rightarrow$ uma maneira possível de determinar política ótima sem enumeração exaustiva (força-bruta).

<u>Representação matricial de uma política</u>:

$$R = \begin{array}{c c} & \begin{array}{cccc} \text{Decisão } 0 & 1 & \dots & K \end{array} \\ \text{Estado} & \\ \begin{array}{c} 0 \\ 1 \\ \vdots \\ M \end{array} & \begin{bmatrix} D_{01} & D_{02} & \dots & D_{0K} \\ D_{11} & D_{12} & \dots & D_{1K} \\ \vdots & \vdots & \vdots & \vdots \\ D_{M1} & D_{M2} & \dots & D_{MK} \end{bmatrix} \end{array}$$

$$D_{ik} = P\left(\text{decisão} = k \middle| \text{estado} = i\right), k = 1,2,\dots,K; i = 0,1,\dots,M$$

com $0 \le D_{ik} \le 1$ e para cada $i \Rightarrow \sum_{k=1}^{K} D_{ik} = 1$

D_{ik}= 0 ou 1$\Rightarrow$ **política determinística**, caso contrário **política aleatória**

<u>Exemplo</u>:

$$R = \begin{array}{c c} & \begin{array}{ccc} \text{Decisão } 1 & 2 & 3 \end{array} \\ \text{Estado} & \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \text{ R determinística}$$

$$R = \begin{array}{c c} & \begin{array}{ccc} \text{Decisão } 1 & 2 & 3 \end{array} \\ \text{Estado} & \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0.4 & 0.6 \end{bmatrix} \end{array} \text{ R aleatória}$$

<u>Probabilidades Conjuntas Estacionárias</u> $\Rightarrow y_{ik}$ = P(estado = i e decisão = k)

conjunta = priori (estacionária)*condicional

priori i = somatória conjuntas variando k

$$\begin{cases} y_{ik} = \pi_i . D_{ik} \\ \pi_i = \sum_{k=1}^{K} y_{ik} \end{cases} \Leftrightarrow D_{ik} = \frac{y_{ik}}{\pi_i} = \frac{y_{ik}}{\sum_{k=1}^{K} y_{ik}}$$

As restrições a que estão sujeitas as probabilidades conjuntas y_{ik} podem ser escritas em função das restrições a que estão sujeitas as probabilidades de estados estáveis π_i, como:

$$\begin{cases} \sum_{i=0}^{M} \pi_i = 1 \Rightarrow \sum_{i=0}^{M}\sum_{k=1}^{K} y_{ik} = 1 \\ \pi_j = \sum_{i=0}^{M} \pi_i . p_{ij} \Rightarrow \sum_{k=1}^{K} y_{jk} = \sum_{i=0}^{M}\sum_{k=1}^{K} y_{ik} . p_{ij}(k) \\ y_{ik} \geq 0, i = 0,1,\ldots,M \quad \text{e} \quad k = 1,2,\ldots,K \end{cases}$$

O custo médio esperado por unidade de tempo fica:

$$E(C) = \sum_{i=0}^{M}\sum_{k=1}^{K} \pi_i C_{ik} D_{ik} = \sum_{i=0}^{M}\sum_{k=1}^{K} C_{ik} . y_{ik}$$

A determinação da Política Ótima torna-se um problema de Programação Linear:

$$\text{Min} \quad \sum_{i=0}^{M}\sum_{k=1}^{K} C_{ik} . y_{ik}$$

$$\begin{cases} \sum_{i=0}^{M}\sum_{k=1}^{K} y_{ik} = 1 \\ \sum_{k=1}^{K} y_{jk} - \sum_{i=0}^{M}\sum_{k=1}^{K} y_{ik} . p_{ij}(k) = 0, \quad \text{para} \quad j = 0,1,\ldots,M \\ y_{ik} \geq 0, i = 0,1,\ldots,M \quad \text{e} \quad k = 1,2,\ldots,K \end{cases}$$

A política encontrada será determinística: para cada i, $y_{ik} > 0$ para pelo menos um k (ao menos uma decisão tem que ser tomada para cada i), o que implica que para cada i, $y_{ik} > 0$ para um único k (existem M+2 restrições e, portanto, M+2 variáveis básicas, sendo que uma restrição é redundante, o que faz com que existam M+1 variáveis com valores # 0), ou seja, D_{ik} = 0 ou 1.

Exemplo: Considerando que quando se toma a decisão 2 no estado 0, a probabilidade do sistema se manter nesse estado aumenta em 0.05 e de passar para o estado 1 diminui em 0.05. As demais probabilidades de transição continuam as mesmas já citadas:

$$P(k=1)= \begin{array}{c|c} \text{Estado} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 0.9 & 0.09 & 0.01 \\ 0 & 0.55 & 0.45 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \quad P(k=2)= \begin{array}{c|c} \text{Estado} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 0.95 & 0.04 & 0.01 \\ 0.8 & 0.2 & 0.0 \\ 0.3 & 0.7 & 0.0 \end{bmatrix} \end{array} \quad P(k=3)= \begin{array}{c|c} \text{Estado} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

Código MPL

custo infinito

```
Min 0Y01 + 50Y02 + 80Y03 + 10Y11 + 50Y12 + 80Y13 + 100000Y21 + 50Y22 + 80Y23

Subject to

Y01 + Y02 + Y03 + Y11 + Y12 + Y13 + Y21 + Y22 + Y23=1;

Y01 + Y02 + Y03 -(0.9Y01 + 0.95Y02 + 1Y03 + 0Y11 + 0.8Y12 + 1Y13 + 0Y21 + 0.3Y22 + 1Y23)=0;

Y11 + Y12 + Y13 -(0.09Y01 + 0.04Y02 + 0Y03 + 0.55Y11 + 0.2Y12 + 0Y13 + 0Y21 + 0.7Y22 + 0Y23)=0;

Y21 + Y22 + Y23 -(0.01Y01 + 0.01Y02 + 0Y03 + 0.45Y11 + 0Y12 + 0Y13 + 1Y21 + 0Y22 + 0Y23)=0;
```

Solução Ótima

MIN Z = 5.7238

Constraint	Slack	Shadow Price
c1	0.0000	5.7238
c2	0.0000	-55.4352
c3	0.0000	0.0000
c4	0.0000	18.9310

1 restrição redundante

Variable	Activity	Reduced Cost
Y01	0.8909	0.0000
Y02	0.0000	47.2327
Y03	0.0000	74.2762
Y11	0.0000	12.7951
Y12	0.1002	0.0000
Y13	0.0000	18.9310
Y21	0.0000	99994.2762
Y22	0.0000	8.7416
Y23	0.0089	0.0000

custo reduzido infinito

Política Ótima

$D_{01} = D_{12} = D_{23} = 1$, demais nulas.