UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA

TRABALHO DE LICENCIATURA

SISTEMA DE BÓNUS MALUS APLICANDO

A PROGRAMAÇÃO DINÂMICA

AUTOR:
PAULINO OFIÇO MACITELE

MAPUTO, MARÇO DE 2005

UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA

TRABALHO DE LICENCIATURA

SISTEMA DE BÓNUS MALUS APLICANDO A PROGRAMAÇÃO DINÂMICA

SUPERVISOR: DR. MÁRIO FRENGUE GETIMANE

AUTOR: PAULINO OFIÇO MACITELE

MAPUTO, MARÇO DE 2005

"Viva como se fosse morrer amanhã;
Aprenda como se fosse viver para sempre"

Mahatma Gandhi (*1869 – 1948*)
Filósofo e Político Indiano

DECLARAÇÃO DE HONRA

Eu, **Paulino Ofiço Macitele**, declaro por minha honra, que este trabalho é de minha autoria e resulta de pesquisa bibliográfica, tratamento e interpretação de dados por mim efectuada. Mais, declaro que, nunca foi apresentado na sua essência, para obtenção de qualquer grau académico.

Maputo, Março de 2006

(Paulino Ofiço Macitele)

DEDICATÓRIA

Aos meus pais, Gonçalo e Laura, por me terem, mostrado o caminho para alcançar algo, nesta caminhada muito complexa.

A minha esposa, Zita pela compreensão e acima de tudo muita amizade durante o curso.

Aos meus irmãos Hofi, Romão, Vinichi, Jossai e Nguize, pela omnipresença em todo meu percurso estudantil.

Por último, ao meu filho do coração, Belito, pela sua existência no mundo.

AGRADECIMENTOS

Ao Dr. Mário Frengue Getimane, supervisor do presente trabalho, pela permanente crítica construtiva, paciência dispensada ao longo da investigação e indicação das fontes de informação.

A todos professores do curso de Licenciatura em Informática, pelos instrumentos teóricos e metodologia de trabalho com que me municiaram ao longo do curso.

Aos meus colegas de turma e de curso, em especial a "*cool*" Gule, Fred, Paulo, Mats, Diogo, Macas, Matsenguas, Bob e outros, Pelo clima de estudo e de camaradagem proporcionados.

Á minha família e amigos pelo encorajamento dado e pela compreensão ao longo da minha carreira estudantil.

Por último, a todos aqueles que de forma directa ou indirecta apoiaram-me na realização deste trabalho.

RESUMO

Neste trabalho, faz-se o estudo de um modelo de programação dinâmica estocástica, em que se considera um acidente, um acontecimento imprevisto ou fortuito do qual resulta um dano à coisa ou à pessoa.

Inicialmente, faz-se a descrição do seguro, como um meio de defesa contra qualquer tipo de risco, até o surgimento dos sistemas de bónus malus bem como as suas características elementares, seguido de alguns exemplos que ajudam melhor a sua compreensão.

Faz-se uma abordagem das principais características e algumas definições das cadeias de Markov que são importantes para com clareza perceber a transição entre as classes do sistema.

Define-se os modelos de programação dinâmica determinística e estocástica e faz-se uma abordagem em termos de acções e de políticas. Apresenta-se alguns teoremas sobre a minimização e maximização dos custos, bem como sobre as propriedades estruturais da solução dum programa dinâmico.

Por fim, faz-se a formulação de um modelo de um sistema de bónus malus, aplicando a programação dinâmica estocástica, com base nas seguintes hipóteses:

- O condutor só pode ter no máximo um acidente por ano,
- Um acidente tem custos aleatórios Y que dependem da gravidade do sinistro,
- Perante um acidente, o condutor pode comunicar à companhia seguradora, neste caso sofrer agravamento no prémio, e pode não comunicar, neste caso concreto obter bonificação no prémio.

Os resultados principais sobre este modelo são que, dependendo de os custos dos acidentes, durante a vigência do contrato serem comparáveis com os prémios a pagar, o condutor pode suportar ele próprio os custos do acidente; caso os custos sejam superiores ao prémio a pagar, então, remeter os acidentes à companhia seguradora. Por outro lado, coloca o problema da necessidade da escala dos prémios ser ajustada aos custos do acidente.

Para terminar, se recomenda a continuação deste trabalho em algumas direcções, que não foi possível abordar neste trabalho, mas que tornam o modelo mais realista, por exemplo, seria interessante, ter em conta que um condutor pode ter durante o ano mais do que um acidente, assim como verificar a dependência do modelo em relação a alguns factores aleatórios como por exemplo a prudência do condutor, a sua rapidez de avaliação e de reacção perante determinadas situações, etc...

GLOSSÁRIO

APÓLICE

É o instrumento do contrato de seguro pelo qual o segurado repassa à seguradora a responsabilidade sobre os riscos, estabelecidos na mesma, que possam advir. A apólice contém as cláusulas e condições gerais, específicas e particulares do contrato e as coberturas especiais e anexos.

DANO

É todo prejuízo material ou pessoal sofrido por um segurado, passível de indemnização, de acordo com as condições de cobertura de uma apólice de seguro.

EVENTO

É toda e qualquer ocorrência ou acontecimento passível de ser garantido por uma apólice de seguro.

INDEMNIZAÇÃO

É a contraprestação do segurador ao segurado que, com a efectivação do risco (ocorrência de evento previsto no contrato), venha a sofrer prejuízos de natureza económica, fazendo jus à indemnização pactuada.

PRÉMIO

É a importância paga pelo segurado à seguradora em troca da transferência do risco a que ele está exposto. Em princípio, o prémio resulta da aplicação de uma percentagem (taxa) à importância da coisa segurada. O prémio deve corresponder ao preço do risco transferido à seguradora.

PROBABILIDADE

Em seguros é a probabilidade de ocorrência de determinado evento coberto pela apólice.

RISCO

É o evento incerto ou de data incerta que não depende da vontade das partes contratantes e contra o qual é feito o seguro. O risco é a expectativa de sinistro. Sem risco não pode haver contrato de seguro. É comum a palavra ser usada, também, para significar a coisa ou pessoas sujeitas ao risco.

SEGURADO

É a pessoa física ou jurídica que, tendo interesses seguráveis, contrata o seguro, em seu benefício pessoal ou de terceiros.

SEGURADORA

É uma instituição que tem o objectivo de indemnizar prejuízos involuntários verificados no património de outrem, ou eventos aleatórios que não trazem necessariamente prejuízos, mediante recebimento de prémios.

SINISTRO

Ocorrência do acontecimento previsto no contrato de seguro e que, legalmente, obriga a seguradora a indemnizar.

TARIFA

Relação das taxas correspondentes a cada classe de risco. É de acordo com a taxa constante da tarifa, que o segurador calcula o prémio relativo ao seguro que lhe é proposto.

TAXA

Elemento necessário à fixação das tarifas de prémios, cálculos de juros, reservas matemáticas, etc. A taxa é uma percentagem fixa dos bens segurados, que se aplica a cada caso determinado, estabelecendo a importância necessária ao fim visado.

ÍNDICE

CAPÍTULO I

1. INTRODUÇÂO

Desde os tempos mais remotos a humanidade tenta encontrar métodos para se proteger contra incertezas do futuro. Mesmo quando as condições de vida eram estritamente ligadas à natureza e os acontecimentos futuros pareciam depender apenas da sorte ou de circunstâncias totalmente alheias à vontade dos homens, as sociedades mais organizadas encontraram meios para atenuar os riscos envolvidos em suas actividades comerciais ou agrícolas.

O medo, a incerteza do amanhã, as ameaças à sua integridade física e ao património sempre acompanhou o homem. Mas o homem inteligentemente concebeu a ideia do seguro, que foi o meio encontrado por si para a sua protecção, prevenindo-o de prejuízos económicos ocasionados pela destruição de seus bens.

O **SEGURO** é um dos meios de defesa contra os riscos de origem natural, humana ou outra, a que se sujeitam os bens e interesses, as actividades do homem e as pessoas.
Exerce importantes funções económicas e sociais, não só porque garante a reposição de patrimónios atingidos por sinistros, contribuindo, assim para o desenvolvimento, mas, também, porque reduz ou atenua a incerteza a um nível aceitável. Seguro é tranquilidade.

Com o desenvolvimento da actividade seguradora, surge nos anos cinquenta, a necessidade de se estabelecer um "preço" justo nos seguros de danos, isto é, para além de tarifar cada segurado de acordo com os valores assumidos por um certo número de variáveis à *priori* que traduzem determinadas características bem definidas, existe necessidade de proceder a revisão à *posterior* tendo em conta a sinistralidade observada do segurado, que é o argumento base dos chamados sistemas de ***bónus malus***.
Os sistemas de ***bónus malus*** são um conjunto de regras de tarifação no seguro automóvel que minoram os prémios futuros das apólices que não participem acidentes à companhia seguradora num determinado período (em geral um ano) e que majoram os prémios futuros das apólices em situação contrária (Ualane,2003).

Estes sistemas foram definitivamente introduzidos em 1960, após que se seguiram vários estudos por parte de Delaporte (1965), Bichsel (1965) e Bühlmann (1964); actualmente são usados em vários países da Europa, Ásia e alguns da América Latina e África.

De referir que em Moçambique, algumas companhias seguradoras usam só sistemas de bonificações em apólices sem sinistros. Por exemplo, se durante uma anuidade de seguro não tiver havido qualquer participação de sinistro, o segurado terá o direito a um desconto de 10% no prémio de renovação; se no decurso de duas ou mais anuidades seguidas não tiver havido qualquer participação de sinistro por desconto passará a ser, nas mesmas condições, adicionado de 10% por cada anuidade até o limite máximo de 40% no fim de 4 anuidades, mantendo-se este quanto as anuidades seguintes, até se verificar qualquer sinistro. A principal diferença entre os sistemas de bónus malus e os sistemas de bonificações em apólices sem sinistros reside no facto de os sistemas de bónus malus majorarem ou minorarem os prémios futuros das apólices que participem ou não acidentes respectivamente; enquanto que os sistemas de bonificações, só descontam os valores do prémio futuro em apólices que não participem qualquer sinistro.

Os sistemas de ***bónus malus*** são importantes, na medida em que permitem detectar, avaliar e introduzir no sistema de tarifação os efeitos conjuntos de factores considerados relevantes para a caracterização do risco transferido para a companhia seguradora, mas que não são directamente mensuráveis como por exemplo a prudência dos condutores, a sua rapidez de avaliação e de reacção perante determinadas situações, etc.... (Ualane,2003).

Com o presente trabalho o estudante pretende conceber, desenvolver e apresentar um modelo de um sistema de ***bónus malus***, usando a programação dinâmica estocástica.

1.1. OBJECTIVO GERAL

Apresentar um modelo de um sistema de bónus malus, com recurso à programação dinâmica estocástica.

1.2. OBJECTIVO ESPECÍFICO

- Definir o sistema de bónus malus (Lemaire, 1998),
- Descrever o funcionamento de um sistema de bónus malus,
- Descrever o modelo determinístico e estocástico de programação dinâmica,
- Propor o modelo, formular o problema, construir o modelo e dar solução de um sistema de bónus malus,
- Discutir a solução,
- Conclusões e recomendações do trabalho.

1.3. MATERIAIS E MÉTODOS

- Para a realização deste trabalho, o estudante recorreu às fontes bibliográficas citadas no trabalho e algumas lições de esclarecimentos de dúvidas por parte do supervisor.

O método descritivo de uma forma geral permitiu:

- Descrever o sistema de bónus malus,
- Exemplificar alguns conceitos,
- Definir alguns conceitos frequentemente usados em seguros,
- Classificar os tipos de programação dinâmica,
- Mostrar os passos essenciais na resolução do problema descrito.

CAPÍTULO II

2. SISTEMAS DE BÓNUS MALUS

Definição* (Lemaire, 1998)*:

Uma companhia de seguros usa um sistema de ***bónus malus*** quando:

- Um grupo de tarifas de segurados pode ser particionado em um número finito de classes, denotadas por C_i ou simplesmente i (i=1,......,s), em que o prémio anual depende somente da classe (s denota o número de classes) e do grupo de tarifas;

- Os novos segurados são agrupados numa classe inicial específica C_{i_0} ;

- Uma classe de segurados num determinado período de seguro (geralmente um ano) é determinada unicamente pelo período precedente e pelo número de sinistros declarados durante o período.

Assim, o sistema é determinado por três elementos:

- A escala de prémios $b=(b_1,.....b_s)$

- A classe inicial (C_{i_0})

- As regras de transição — as regras que determinam a transferência de uma classe para outra quando o número de sinistros é conhecido.

Estas regras podem ser apresentadas como transformações T_k, tais que $T_k(i)=j$ se a apólice de seguros é transferida da classe C_i para a classe C_j quando k sinistros tiverem sido reportados. O termo T_k pode ser escrito sob forma matricial

$$T_k = (t_{ij}^{(k)}),$$

onde $t_{ij}^{(k)} = 1$ se $T_k(i) = j$ e 0 no caso contrário. A probabilidade $P_{ij}(\lambda)$ de a apólice mover-se de C_i para C_j num período, para um segurado caracterizado por alguns parâmetros λ (pela circunstância, frequência de sinistros), é igual a:

$$p_{ij}(\lambda) = \sum_{k=0}^{\infty} p_k(\lambda) t_{ij}^{(k)},$$

onde $P_k(\lambda)$ é a probabilidade de um condutor com uma frequência de sinistros λ ter k sinistros por ano. Obviamente $P_{ij}(\lambda) \geq 0$ e $\sum_{j=1}^{s} p_{ij}(\lambda) = 1$.

A matriz $M(\lambda) = [P_{ij}(\lambda)] = \sum_{k=0}^{\infty} p_k(\lambda) T_k$, é a matriz de transição desta cadeia de Markov.

Um segurado entra no sistema na classe inicial quando obtém a sua licença de condução. Então, durante todo o tempo de vida de condutor, as regras de transição são aplicadas em cada renovação para determinar a nova classe em função do historial dos sinistros.

Na definição anterior, o sistema de bónus malus forma uma cadeia de Markov. Uma cadeia da primeira ordem de Markov é um processo estocástico no qual o desenvolvimento futuro depende somente do estado presente mas não do historial do processo ou da maneira como este foi alcançado.

Toma-se como exemplo de um sistema de bónus malus o do Brasil; um sistema em que os segurados são subdivididos em apenas sete classes, com níveis do prémio 100, 90, 85, 80, 75, 70 e 65, os novos segurados têm que começar na classe 7, no nível 100. Cada ano sem reportar um sinistro, resulta em um desconto de uma classe, e a cada sinistro reportado é penalizado em uma classe. Por outro lado, quando um condutor atinge a classe 1, que é a classe máxima de descontos, ele poderá permanecer nesta classe no caso em que não declare qualquer sinistro, mas no caso de declarar qualquer sinistro, terá como resultado a penalização para uma classe, conforme as regras de transição.

As regras de transição são apresentadas na tabela seguinte:

Classes	Prémio	Classes seguintes						
		0 sinistro	1 sinistro	2 sinistros	3 sinistros	4 sinistros	5 sinistros	≥ 6 sinistros
7	100	6	7	7	7	7	7	7
6	90	5	7	7	7	7	7	7
5	85	4	6	7	7	7	7	7
4	80	3	5	6	7	7	7	7
3	75	2	4	5	6	7	7	7
2	70	1	3	4	5	6	7	7
1	65	1	2	3	4	5	6	7

Exemplos

1. Um segurado entra no sistema na classe 7, equivalente ao nível de prémio 100. Caso não participe nenhum sinistro durante a anuidade do seguro, sofrerá um desconto de uma classe, neste caso passará para a classe 6, equivalente ao nível de prémio 90, todavia, se durante a anuidade do seguro participar algum sinistro, ele se manterá na classe 7, equivalente ao nível de prémio 100. Supondo que o segurado em cada anuidade do seguro não participe qualquer sinistro, será descontado uma classe por cada anuidade até que atinja a classe 1, equivalente ao nível de prémio 65.
2. Consideremos que um segurado se encontra na classe 3, equivalente ao nível de prémio 75. Se na anuidade do seguro, não participar nenhum sinistro, sofrerá o desconto de uma classe, isto é, transita para a classe 2, equivalente ao nível de prémio 70; caso participe 1 sinistro, transitará para a classe 4, equivalente ao nível de prémio 80; caso participe 2 sinistros, transitará para a classe 5, equivalente ao nível de prémio 85; caso participe 3 sinistros, transitará para a classe 6, equivalente ao nível de prémio 90; caso participe pelo menos 4 sinistros, transitará para a classe 7, equivalente ao nível de prémio 100 e dependendo da sua evolução, poderá ser bonificado ou penalizado.

Em geral, na implementação dos sistemas de bónus malus, cada país é livre de usar a escala que lhe convier assim como cada companhia de seguros é livre de aplicar a sua escala dependendo das regras, regulamentos e ambiente adequados, isto é, há países que têm uma escala única, há outros que permitem a variação entre empresas.

Vamos através das regras de transição do sistema brasileiro, mostrar exemplos das matrizes $T_k = (t_{ij}^{(k)})$. No geral esta matriz pode ser escrita como segue:

$$T_k = t_{ij}^{(k)} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} & t_{15} & t_{16} & t_{17} \\ t_{21} & t_{22} & t_{23} & t_{24} & t_{25} & t_{26} & t_{27} \\ t_{31} & t_{32} & t_{33} & t_{34} & t_{35} & t_{36} & t_{37} \\ t_{41} & t_{42} & t_{43} & t_{44} & t_{45} & t_{46} & t_{47} \\ t_{51} & t_{52} & t_{53} & t_{54} & t_{55} & t_{56} & t_{57} \\ t_{61} & t_{62} & t_{63} & t_{64} & t_{65} & t_{66} & t_{67} \\ t_{71} & t_{72} & t_{73} & t_{74} & t_{75} & t_{76} & t_{77} \end{bmatrix}$$

Para o caso em que o condutor, apresenta $k=0$ sinistros, teremos a seguinte matriz T_k:

$$T_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

Analogamente, se o condutor, apresentar $k=4$ sinistros, teremos a seguinte matriz T_k:

$$T_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

CAPÍTULO III

3. CADEIAS DE MARKOV

As cadeias de Markov proporcionam modelos matemáticos para uma variedade de sistemas que ao longo do tempo, se movimentam através de um conjunto enumerável de estados. A evolução desses sistemas é regida por um mecanismo probabilístico cuja característica principal está na forma de dependência entre o comportamento passado e futuro, do sistema, relativamente a um instante fixo. Especificamente, o comportamento futuro do sistema a partir de um instante fixo depende de seu comportamento passado apenas através do estado que o sistema ocupa naquele instante.

***Definição*:**

Um processo aleatório com parâmetro t, $X(t)$, é um processo de Markov se o futuro do processo dado o presente é independente do passado. Isto é, para os instantes arbitrários,

$t_1 < t_2 < \ldots < t_k < t_{k+1}$:

$P[x(t_{k+1}) = x_{k+1}/x(t_k) = x_k, \ldots x(t_1) = x_1] = P[x(t_{k+1}) = x(t_{k+1})/x(t_k) = x_k]$, se $x(t)$ assumir os valores discretos;

$P[a < x(t_{k+1}) < b/x(t_k) = x_k, \ldots, x(t_1) = x_1] = P[a < x(t_{k+1}) < b/x(t_k) = x_k]$,se $x(t)$ assumir os valores contínuos.

***Definição*:**

Uma cadeia de Markov em tempo discreto $\{X_n\}$ é um processo de Markov com espaço de parâmetros discreto e espaço de estados discreto.

***Definição*:**

São designadas probabilidades de transição num passo, a probabilidade do processo estar no estado j, no instante $n+1$, se estava no estado i no instante n e é representado por:

$$P_{i,j}^{n,n+1} = \text{Prob}\ \{ X_{n+1} = j \backslash X_n = i \}.$$

Definição:

Diz-se que uma cadeia de Markov é homogénea (com probabilidade de transição estacionária) se as probabilidades de transição num passo, forem independentes do tempo *n*, isto é, $p_{ij}^{n,n+1} = p_{ij}$ que não depende do tempo *n*.

As probabilidades de transição podem ser representadas matricialmente:

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots p_{0n} \\ p_{10} & p_{11} & p_{10} & \cdots p_{1n} \\ \cdots & \cdots & \cdots\cdot & \cdots \\ p_{n0} & p_{n1} & p_{n2} & \cdots p_{nn} \end{bmatrix};$$

Onde $\{0, 1, 2, \ldots, n\}$ é o conjunto dos estados do sistema.

A matriz *P*, é chamada matriz de Markov ou matriz de probabilidades de transição do processo e goza das seguintes propriedades:

$$p_{ij} \geq 0; i, j = 0, 1, 2, \ldots$$

$$\sum_{j=0}^{n} p_{ij} = 1$$

A matriz de probabilidades de transição em *n* passos é a matriz, $P^{(n)} = [p_{ij}^{(n)}]$, onde $[p_{ij}^{(n)}]$é a probabilidade do processo passar do estado *i* para *j*, em *n* passos.

Pode-se mostrar que

$$P^{(n)} = P^n = \underbrace{P.P.P.,\ldots,.P}_{n \text{ vezes}} \; .$$

Exemplo

Um sistema S qualquer, pode se encontrar nos estados A_1, A_2 *e* A_3. As transições de um estado para o outro estado ocorrem em concordância com o esquema da cadeia de Markov homogénea; as probabilidades de transição são dadas pela matriz:

$$P = \begin{bmatrix} 1/2 & 1/6 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Esta matriz diz-nos que, se o sistema se encontra no estado A_1, após um passo, ele ficará no mesmo estado com a probabilidade de 1/2, e transitará para o estado A_2 com a probabilidade de 1/6, e para o estado A_3 com a probabilidade 1/3. Contudo, se o sistema estivesse no estado A_2, então após um passo, poderia (com igual probabilidade) encontrar-se somente nos estados A_1 e A_3, pois não pode transitar do estado A_2 para A_2. A ultima linha da matriz mostra-nos que do estado A_3 o sistema pode transitar para qualquer estado possível com uma e mesma probabilidade de 1/3 (Gnedenko, 1978).

CAPÍTULO IV

4.1. MODELO DETERMINÍSTICO DE PROGRAMAÇÃO DINÂMICA

O modelo determinístico de programação dinâmica é usado em problemas de decisão determinística onde é possível prever o futuro estado do sistema sendo conhecido o estado inicial e a sequência das decisões.

A exposição de programação dinâmica que se segue, guia-se por Hinderer (1991).

4.1.1 DEFINIÇÃO BÁSICA

Definição :

Consideremos os conjuntos B e C não vazios e seja $M \subset B \times C$.

O conjunto $M(b) = \{c:(b,c) \in M\}$ chama-se *b-secção* de *M*.

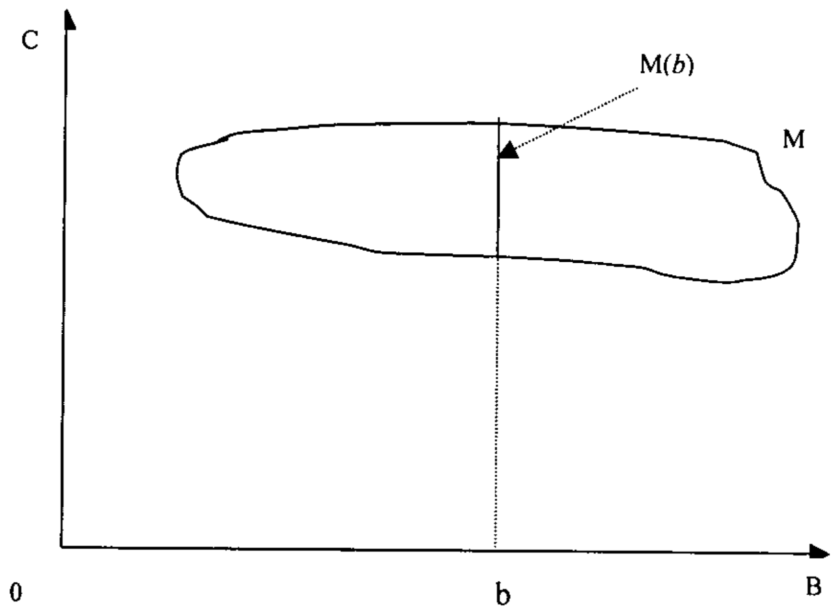

Fig. 1 : *M*(*b*) é secção de *M*

Definição :

Um ***modelo determinístico*** de programação dinâmica, é um vector $(S, A, D, T, r, V_0, \beta)$ onde:

a) S é um conjunto não vazio, chamado ***espaço de estados***.

b) A é um conjunto não vazio, chamado ***espaço de acções***.

c) D é um subconjunto de $S \times A$ tal que todas as secções $D(s)$ de D não são vazias.
D é chamado ***conjunto restrição*** e $D(s)$ é o conjunto de ***acções admissíveis para o estado s.***

d) T é uma função de D em S, chamada ***função de transição***.

e) r é uma função finita em D, chamada ***função compensatória*** num estágio.

f) V_0 é uma função finita em S, chamada ***função terminal compensatória***.

g) β é um número real positivo, chamado ***factor de desconto***.

Mediante qualquer problema, o primeiro passo para encontrar a solução é modelá-lo através da escolha de S, A, D, T, r, V_0 e β, onde o mais difícil é a escolha do espaço S.

4.1.2 FORMULAÇÃO DO PROBLEMA DE OPTIMIZAÇÃO DO MODELO DE PROGRAMAÇÃO DINÂMICA

4.1.2.1. FORMULAÇÃO EM TERMOS DE ACÇÕES

Definição :

Dado um estado inicial arbitrário s_0, afirmaremos que a sequência de acções $(a_0, a_1, ..., a_{N-1}) \in A^N$ é admissível para s_0, se

$$a_o \in D(s_0),$$
$$a_1 \in D(s_1), \text{ onde } s_1 := T(s_0, a_0),$$
$$a_2 \in D(s_2), \text{ onde } s_2 := T(s_1, a_1),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{N-1} \in D(s_{N-1}), \text{ onde } s_{N-1} := T(s_{N-2}, a_{N-2})$$

Colocando $s_N := T(s_{N-1}, a_{N-1})$, a sequência $(s_1, s_2, ..., s_N)$ é chamada ***sequência de estados gerada pelo estado inicial s_0 e pela sequência de acções admissíveis*** $y := (a_n)_0^{N-1}$.

Note-se que s_n é uma função de s_0 e y: $s_n = s_{ny}(s_0)$.

O conjunto de sequências de acções admissíveis para o estado inicial s_0 será denotado por $A^N(s_0)$. Em particular, se $D(s) = A$ para qualquer s, então $A^N(s_0) = A^N$ para todo s_0.

A equação:

$$s_{n+1} = T(s_n, a_n), \; 0 \le n \le N-1,$$

que para quaisquer que sejam s_0 e $(a_n)_0^{N-1} \in A^N(s_0)$ dados , determina a evolução do sistema, é chamada *equação de estado* ou *equação do sistema.*

Para cada estado inicial $s_0 \in S$ e cada sequência de acções $y := (a_n)_0^{N-1} \in A^N(s_0)$, o *ganho* ao fim de N períodos é definido como o número real

$$V_{Ny}(s_0) := \sum_{n:=0}^{N-1} \beta^n r(s_n, a_n) + \beta^N V_0(s_N),$$

onde $(s_n)_1^N = (s_{ny}(s_0))_1^N$ é uma sequência de estados gerados por s_0 e y.

A primeira formulação do problema de optimização é a seguinte:

I. Calcular a recompensa máxima ao fim de N períodos, sup $\{V_{Ny}(s_0) : y \in A^N(s_0)\}$, para o sistema começando em s_0.

II. Encontrar, se possível, uma sequência de acções óptimas geradas por s_0, isto é o ponto máximo da função $y \to V_{Ny}(s_0)$.

4.1.2.2. FORMULAÇÃO EM TERMOS DE POLÍTICAS

Esta formulação é particularmente útil para uma apresentação sucinta do procedimento da solução básica e que se torna indispensável quando deparamos com um modelo de programação dinâmica estocástico.

Uma política pode ser vista como um plano que especifica a decisão a ser tomada em cada estágio para cada estado que possa ocorrer nesse estágio.

Política óptima é uma política que maximiza (ou minimiza) o *ganho* (ou *custo*) ao fim dum certo número de estágios.

Definição :

I. A função $\Phi: S \to A$ tal que $\Phi(s) \in D(s)$ para todo $s \in S$ é chamada ***regra de decisão***.

Uma sequência $\pi = (\Phi_0, \Phi_1, \ldots, \Phi_{N-1})$ de N regras de decisão é chamada ***política de N períodos***.

O conjunto não vazio de regras de decisão será denotado por F; F^N é o conjunto de todas as políticas de N períodos; ele não depende do estado inicial. A condição $\Phi(s) \in D(s)$ para a regra de decisão Φ, que é equivalente a gráfico de $\Phi \subset D$, garante que a decisão tomada, quando a política de N períodos é usada, selecciona para cada estágio e cada estado somente acções admissíveis.

Se o sistema começar no estado s_0 e aplicamos a política $\pi = (\Phi_n)_0^{N-1} \in F^N$, então este move-se sucessivamente através da sequência de estados

$$s_1 := T(s_0, \Phi_0(s_0))$$

$$s_2 := T(s_1, \Phi_1(s_1))$$

....

$$s_N := T(s_{N-1}, \Phi_{N-1}(s_{N-1}))$$

Note que s_n, $1 \le n \le N$, depende de s_0 e de $\pi : s_n = s_{n\pi}(s_0)$. Assim, a equação do sistema tem a forma

$$s_{n+1} = T(s_n, \Phi_n(s_n)), \qquad 0 \le n \le N-1.$$

Chamamos $(s_n)_1^N$ ***a sequência de estados gerada por s_0 e pela política*** π e $(\Phi_n(s_n))_0^{N-1}$ ***a sequência de acções gerada por s_0 e π***; obviamente (s_n) é também a sequência de estados gerada por s_0 e pela sequência de acções $(\Phi_n(s_n))_0^{N-1}$. Associada ao desenvolvimento do sistema a partir do estado inicial s_0, está a recompensa ou ganho total ao fim de N períodos sob a política π definida por

$$V_{N\pi}(s_0) := \sum_{n:=0}^{N-1} \beta^n r(s_n, a_n) + \beta^N V_0(s_N).$$

A segunda formulação do problema de optimização é a seguinte:

I. Calcular a recompensa máxima ao fim de N períodos, isto é, $\sup\left\{ V_{N\pi}(s_0) : \pi \in F^N \right\}$ para o estado inicial s_0.

II. Encontrar, se possível, uma política óptima começando no estado s_0, isto é, o ponto máximo da função $\pi \to V_{N\pi}(s_0)$.

III. Encontrar, se possível, uma ***política óptima***, para cada estado inicial s_0.

4.1.2.3. RELAÇÃO ENTRE AS DUAS FORMULAÇÕES

Proposição : Seja s_0 um estado inicial arbitrário

I. É válida a seguinte igualdade

$$\sup\left\{V_{Ny}(s_0): y \in A^N(s_0)\right\} = \sup\left\{V_{N\pi}(s_0): \pi \in F^N\right\} = V_N(s_0).$$

II. Se π^* é uma política de N estágios e se y^* é uma sequência de acções geradas por s_0 e pela política π^*, então π^* é óptima para s_0 sse y^* é óptima em relação a s_0.

Nota: O número $V_N(s_0)$ – que pode ser igual a $+\infty$, é a recompensa máxima ao fim de N períodos. A função $V_N : S \to R$ é chamada ***função do valor*** do modelo de programação dinâmica de N estágios. É de salientar que $V_N(s_0)$ é finita se existe uma sequência de acções óptima em relação a s_0.

<u>*Demonstração*</u>:

Para cada $\pi \in F^N$ passaremos a denotar por $f(\pi)$ a sequência de acções geradas por s_0 e π. Seja $y := (a_n)_0^{N-1}$ uma sequência de acções arbitrárias admissíveis para s_0 e seja $(s_n)_1^N$ a sequência de estados gerados por s_0 e y. Seja

$$V_{Ny}(s_0) := \sum_{n:=0}^{N-1} \beta^n r(s_n, a_n) + \beta^N V_0(s_N) \text{ e } V_{N\pi}(s_0) := \sum_{n:=0}^{N-1} \beta^n r(s_n, a_n) + \beta^N V_0(s_N).$$

Para provar que $V_{Ny}(s_0) = V_{N\pi}(s_0)$ é suficiente mostrar que $r(s_n, a_n) = r(s_n, \Phi_n(s_n))$. Colocando, para $0 \le n \le N-1$,

$$\phi_n(s) := \begin{cases} a_n, se\ s = s_n \\ arbitrário\ em\ D(s),\ para\ outros\ casos \end{cases}$$

é claro que, $y = f(\pi)$ para $\pi := (\phi_n)_0^{N-1}$.

4.1.2.4. MÉRITOS E DEFEITOS DE CADA FORMULAÇÃO

- Problemas teóricos são mais fáceis de formular em termos de políticas do que em termos de sequência de acções. Por outro lado, a solução de um problema prático é geralmente formulada em termos de sequência de acções.

- Maximizar sobre um conjunto de sequências de acções, à primeira vista parece mais fácil que maximizar sobre um conjunto de políticas. Esta impressão é reforçada pelo facto de que em geral o conjunto F^N de políticas de N estágios possuí uma cardinalidade maior que os conjuntos $A^N(s_0)$. Consequentemente, existe uma maior abundância de políticas óptimas em casos onde há relativamente poucas sequências de acções óptimas.

- Em modelos de programação dinâmica estocásticos os estados momentâneos e acções tomadas em geral não determinam dum modo único o futuro. Neste, caso, a formulação em termos de politicas é apropriada.

4.1.3. MÉTODO DE SOLUÇÃO RECURSIVA BÁSICA

Vamos de seguida fazer a formulação do método fundamental de solução do problema de optimização dum programa dinâmico de N períodos.

Lema : Sejam S e A conjuntos não vazios e v uma função definida no conjunto $D \subset S \times A$ tal que $D(s) \neq \phi$, $\forall\ s \in S$. então

$$\sup_{(s,a)\in D} v(s,a) = \sup_{s\in S} \sup_{a\in D(s)} v(s,a)$$

Em particular, se $D = S \times A$, então

$$\sup_{(s,a)\in D} v(s,a) = \sup_{s\in S} \sup_{a\in A} v(s,a) = \sup_{a\in A} \sup_{s\in S} v(s,a)$$

Vamos usar o lema descrito na página 17, para uma abordagem recursiva dum problema num certo estágio n, para $n \geq 2$. Usando a formulação em termos de acções, o método será descrito do seguinte modo:

Assume-se que é conhecida a recompensa máxima no estágio $n-1$, $V_{n-1}(t)$, para qualquer estado inicial t.

O ganho máximo esperado $V_n(s)$ para o modelo no estágio n com o estado inicial s_0, é um número real e pode ser calculado recursivamente através da ***iteração de valor***:

$$V_n(s) = \max \{ W_n(s,a) : a \in \mathrm{D}(s),\ n \in N,\ s \in S \}$$

onde

$$w_n := r(s,a) + \beta V_{n-1}(\mathrm{T}(s,a)),$$

Portanto, a recompensa $V_n(s)$ será obtida, maximizando a função $w_n(s,.)$.

Consideremos um modelo de programação dinâmica. Se o sistema está no estado s num certo instante n e aplicamos uma regra de decisão $f \in F$, então a função r_f será definida do seguinte modo: $r_f(s) := r(s,f(s))$.

Se f é uma regra de decisão e $\sigma = (\mu_1, \mu_2, \ldots, \mu_{n-1})$ é uma política de $n-1$ estágios, então (f, σ) é uma política de n estágios.

De referir que em programação dinâmica, há distinção entre estágio e período, pois os períodos contam-se de trás para frente e os estágios de maneira contrária, assim, quando se fala de estágio 1, significa que falta um período para chegar ao fim, ou seja estamos no período $N-1$.

4.1.4. TEOREMA BÁSICO PARA UM MODELO ESTACIONÁRIO DETERMINÍSTICO DE PROGRAMAÇÃO DINÂMICA

O teorema básico para um modelo determinístico, é usado para o problema de maximização de ganhos, e é enunciado como segue:

Teorema :

a) É válida a ***iteração de valor***:

$$V_n(s) = \sup_{a \in D(s)} \{ r(s,a) + \beta V_{n-1}(T(s,a)) \} = \sup_{f \in F} \{ r_f(s) + \beta V_{n-1}(T_f(s)) \},$$

onde $s \in S$ e $n \in N$.

b) É válido o seguinte ***critério de optimalidade***: Se $f_n(s)$ é um ponto máximo da função

$$a \to w_n(s,a) := r(s,a) + \beta V_{n-1}(T(s,a)), \text{ para } 0 \le n \le N \text{ e } s \in S,$$

então a politica de N períodos $\pi^* := (f_n)_N^1$ é óptima. Além disso, para $s_0^* = s_0 \in S$, se $(s_n^*)_1^N$ é gerada por s_0 e pela politica π^*, então a sequência de acções

$a_n^* := f_{N-n}(s_n^*)$, $0 \le n \le N-1$, é óptima para s_0.

Definição: O conjunto $D_n^*(s)$ dos pontos máximos da função $a \to w_n(s,a)$ é chamado conjunto de ***acções óptimas***.

A regra de decisão f é um ***maximizador*** no estágio n se $f(s) \in D^*(s)$ para todo $s \in S$.

Nota: A regra de decisão f é um maximizador no estágio n se $W_n(s,f(s)) = V_n(s)$.

Com a definição anterior o critério de optimalidade poderá ser enunciado duma forma mais curta como se segue:

Se, para $1 \le n \le N$, f_n é um maximizador no estágio n, então a sequência $(f_n)_N^1$ é óptima.

É importante salientar que uma politica óptima não precisa de ser composta por maximizadores. Neste contexto, se $\pi = (f_n)_N^1$ é uma sequência óptima, então $f_n(s)$ não é necessariamente elemento do conjunto $D_n^*(s)$.

4.1.5. TEOREMA BÁSICO PARA MINIMIZAÇÃO DE CUSTOS

Para o problema de minimização dos custos, o teorema básico é enunciado como segue:

a) É válida a ***iteração de valor***:

$$C_n(s) = \inf_{a \in D(s)} \{ -r(s,a) + \beta C_{n-1}(T(s,a)) \} = \inf_{f \in F} \{ -r_f(s) + \beta C_{n-1}(T_f(s)) \},$$

onde $s \in S$ e $n \in N$.

b) É válido o seguinte ***critério de optimalidade***: se para $1 \le n \le N$, f_n é minimizador no estágio n, então a política $\pi^* := (f_n)_N^1$ é óptima.

Exemplo (Hinderer, 1991)

Alocação entre consumo e reinvestimento:

Uma companhia promovendo um projecto, dispõe de $K \in R^+$ unidades monetárias. No inicio de cada período de tempo N, $n = 0, 1, 2, \ldots, N-1$ a companhia deve alocar o capital monetário entre as seguintes possibilidades: Consumir (para a directa promoção do projecto) ou reinvestir. Assume-se que o consumo de $b \in R_+$ unidades, resulta, numa utilidade numérica $u(b) \in R_+$ e que $i \in R^+$ é a taxa de lucro por período de reinvestimento. Assim, $\gamma := 1 + i$ é o factor de lucro. Como poderá o capital ser alocado no inicio de cada período de tempo N, no sentido de maximizar a soma da utilidade num estágio? Quão grande é esta utilidade total máxima?

Solução:

Vamos dar uma solução detalhada para o caso especial em que a utilidade num estágio é de $u(b) := \sqrt{b}$, que é a recompensa terminal da formula $V_0(s) = d_0\sqrt{s}$ para qualquer $d_0 \in R_+$ e que o factor de desconto β para o consumo é arbitrário. Os mais interessantes parecem ser os dois (2) casos onde o capital final não consumido, é retirado ($d_0 = 0$) ou valorizado como o consumo ($d_0 > 0$). (note-se também que no caso geral, $u(b) := c\sqrt{b}$ para qualquer $c \in R^+$ é facilmente reduzido para o presente caso, substituindo d_0 por d_0/c.)

Primeiro resolve-se o problema de um estágio (para o primeiro estágio) usando

$a :=$ proporção da quantidade do capital consumido

$$V_1(s) = \sup_{0 \le a \le 1} \left[\sqrt{as} + \beta V_0(T(s,a))\right] = \sup_{a} \left[\sqrt{as} + \beta d_0 \sqrt{\gamma s(1-a)}\right]$$

$$= \sqrt{s}\,.\, \sup_{a} \left[\sqrt{a} + c\sqrt{(1-a)}\right];$$

onde colocamos $c := \beta d_0 \sqrt{\gamma}$ de modo que $c \in R_+$. Se $c = 0$ então

$a \to g(a) := \sqrt{a} + c\sqrt{1-a}$, $a \in [0,1]$, tem como ponto máximo $a^* = 1$. E se $c > 0$ então,

$g'(a) = 1/(2\sqrt{a}) - c/(2\sqrt{1-a})$, $a \in (0,1)$.

Obviamente, g' é decrescente e $g'(a^*) = 0$ se $a^* = 1/(1 + c^2)$; note que $a^* \in (0,1)$. Dado que g é contínuo no intervalo [0,1], sabemos que a^* é o ponto máximo de g. Segue que para $c \geq 0$,

$S \to f_1(s) = 1/(1 + c^2) = 1/(1 + \beta^2 d_o^2 \gamma)$ é o maximizador no estágio $n = 1$. Além disso, obteremos facilmente de

$V_1(s) = \sqrt{s}\,.\,g(a^*) = \sqrt{1+c^2}\sqrt{s}$ que, $V_1(s) = \sqrt{1 + \beta^2 \gamma d^2_o}\,.\,\sqrt{s}$.

A solução do problema de um estágio, mostra duas características notáveis:

(i) O maximizador f_1 não depende do estado. Esta é uma situação muito rara. Note que teríamos obtido um maximizador $f_1(s) = 1/(1 + \beta^2 d_o^2 \gamma)$. s neste caso, foi definido a_n não para ser a proporção, mas sim a quantidade do capital consumido no tempo n.

(ii) O valor da função V_1 dum estágio é da mesma "formula funcional" como V_0 nomeadamente $V_1(s) = d_1\sqrt{s}$, onde $d_1 \in R_+$. Consequentemente, substituindo d_1 e d_0 acima, obteremos $V_2(s) = d_2\sqrt{s}$ para qualquer $d_2 \in R_+$ e um maximizador constante f_1 etc. Mais formalmente por indução em n e $V_1(s) = \sqrt{1 + \beta^2 \gamma d^2_o}\,.\,\sqrt{s}$, obteremos que:

(i) V_n está sob forma de

$V_n(s) = d_n\sqrt{s}$, $s \in R^+$, para qualquer $d_n \in R_+$, que satisfaz a fórmula recursiva

$d_n = \sqrt{1 + \beta^2 \gamma d^2_{n-1}}$, $n \in N$.

(ii) A regra de decisão

$f_n \equiv 1/d^2_n$ é um maximizador no estágio n, donde vem que $(f_n)^1_N$ é uma política óptima para N estágio. A equação recursiva descrita na página anterior, é facilmente resolvida e dá: $d_n^2 = \sum_{\nu=0}^{n-1} (\beta^2\gamma)^\nu + d_o^2(\beta^2\gamma)^n$.

As somas tais como a descrita acima, quando ocorrem, serão abreviadas da seguinte forma:

$$\sigma_n(x) := \sum_{\gamma=0}^{n-1} x^\nu = \begin{cases} (1-x^n)/(1-x), \; x \in R - \{1\}, \\ n, \; x = 1. \end{cases}$$

A equação recursiva $d_n = \sqrt{1 + \beta^2\gamma d^2_{n-1}}$, $n \in N$ ocorre frequentemente e a sua solução pode ser escrita como segue:

Lema: (solução linear da equação da diferença de 1ª ordem)

Se a sequência $(b_n)_0^\infty$ de números reais satisfaz

$b_n = c + \alpha b_{n-1}$, $n \in N$, para qualquer c e α reais, então segue que

$b_n = c \,.\, \sigma_n(\alpha) + \alpha^n \,.\, b_0$, $n \in N$.

Assim, a expressão $d_n^2 = \sum_{\nu=0}^{n-1} (\beta^2\gamma)^\nu + d_o^2(\beta^2\gamma)^n$, assume a forma:

$$d^2_n = \sigma_n(\beta^2\gamma) + d_o^2(\beta^2\gamma)^n.$$

As equações, $V_n(s) = d_n\sqrt{s}$, $f_n \equiv 1/d^2_n$ e $d_n^2 = \sum_{\nu=0}^{n-1} (\beta^2\gamma)^\nu + d_o^2(\beta^2\gamma)^n$, constituem a solução do problema. como $f_n(s)$ não depende de s, a computação de uma acção s_0 óptima da sequencia $(a_n^\bullet)_0^{N-1}$, não requer os procedimentos anteriores, assim teremos a quantidade do capital consumido no tempo n como segue: $a_n^\bullet = f_{N-n} = 1/d^2_{N-n}$.

4.2. MODELO ESTOCÁSTICO DE PROGRAMAÇÃO DINÂMICA

O modelo estocástico de programação dinâmica, é usado para situações em que o resultado de cada decisão não é completamente previsível, mas pode ser observado antes que a próxima decisão seja tomada.

O caso mais simples de um modelo estocástico de N períodos é aquele em que a transição de s_n para s_{n+1} é especificada pela função de transição T a qual é "perturbada" por alguma variável aleatória X_{n+1} tomando valores num conjunto finito M, e onde as variáveis X_i's são independentes e identicamente distribuídas (i.i.d.). Em muitas aplicações deste modelo o espaço de estados S será contável ou mesmo finito.

A lei de transição para este modelo estocástico é uma função T de $D \times M$ em S com o seguinte significado intuitivo: se num instante n estamos no estado s_n e tomamos a acção a_n e se a disturbância X_{n+1} toma o valor x_{n+1} então o sistema move-se para o novo estado aleatório $s_{n+1} :=$ T(s_n, a_n, x_{n+1}), $0 \leq n \leq N-1$; esta equação deve ser entendida do seguinte modo: dada uma política de N períodos $\pi = (\varphi_n)_0^{N-1}$ e um estado inicial s_0, o próximo estado do sistema é uma variável aleatória $\zeta_{n\pi}$ definida recursivamente pela equação estocástica $\zeta_{n+1,\pi} = T_{\varphi_n}(\zeta_{n\pi}, X_{n+1})$, $0 \leq n \leq N-1$,

onde $T_{\varphi_n}(\zeta_{n\pi}, X_{n+1}) = T(\zeta_{n\pi}, \varphi_n(\zeta_{n\pi}), X_{n+1})$, $\varphi \in F$ com a condição inicial $\zeta_{0\pi} \equiv s$.

Note-se que:

$\zeta_{n\pi}$ é uma função do vector aleatório $Y := (X_i)_1^N$, mas não depende de $X_{n+1}, \ldots, X_N$. Portanto $\zeta_{n\pi}$ e X_{n+1} são estocasticamente independentes.

Para s_0 e π fixos, $\zeta_{n\pi}$ apenas assume um número finito de valores, já que M é finito.

4.2.1. DEFINIÇÃO BÁSICA

O modelo estocástico de programação dinâmica é um vector (S, A, D, M, q, T, r, V_0, β) com as seguintes características:

a) S, A, D, r, V_0, β têm o mesmo significado e interpretação como no modelo determinístico de programação dinâmica.

b) M é um conjunto finito cujos elementos são *valores de disturbância*

c) q é uma função de *densidade de probabilidade das disturbância i.i.d.*

d) T é uma função de $D \times M$ em S, chamada *função de transição estocástica.*

A recompensa ao fim de N estágios, quando o sistema começa em s_0 é a variável aleatória real

$$R_{N\pi}(s,Y) := \sum_{n=0}^{N-1} \beta^n r(\zeta_{n\pi}, \varphi_n(\zeta_{n\pi})) + \beta^N V_0(\zeta_{N\pi}), (**)$$

onde $Y := (X_i)_1^N$ e $V_0(\zeta_{N\pi})$ é o ganho terminal no fim do processo.

Além disso, o ganho esperado ao fim de N períodos é definido pelo número real

$V_{N\pi}(s) := E[R_{N\pi}(s,Y)] = \sum_{y \in M^N} R_{N\pi}(s,y).P(Y = y)$, para o horizonte N. A *função de valor*

$$V_N : S \to (-\infty, \infty] \text{ é definida por } s \to V_N(s) := \sup\{ V_{N\pi}(s) : \pi \in F^N,\ N \in \mathbb{N} \}.$$

Lema : O ganho esperado ao fim de N etapas $V_{N\pi}$ pode ser obtido iterativamente para cada horizonte N e cada política π do seguinte modo:

$$V_{n(f,\sigma)}(s) = r_f(s) + \beta E[V_{n-1,\sigma}(T_f(s,X))],\ n \geq 2, (f,\sigma) \in F \times F^{n-1},$$

$$V_{1f}(s) = r_f(s) + \beta E[V_0(T_f(s,X))],\ f \in F,\ s \in S.$$

<u>Demonstração</u>: consideremos o caso $n \geq 2$, para $s \in$ S fixo e $\pi := (f, \sigma)$, onde

$\sigma = (\varphi_1, \varphi_2, \ldots, \varphi_{N-1})$, $Y_2 := (X_2, X_3, \ldots, X_n)$, $y_2 := (x_2, x_3, \ldots, x_n)$, então,

$P(Y = (x_1, y_2)) = P(Y_2 = y_2).P(X_1 = x_1)$ uma vez que as disturbâncias X_v são i.i.d.

A partir de (**) obtemos $R_{n\pi}(s, Y) = r_f(s) + \beta R_{n-1,\sigma}(T_f((s, X_1), Y_2))$. Segue-se que

$V_{n\pi}(s) = r_f(s) + \beta E[R_{n-1,\sigma}(T_f(s, X_1), Y_2)]$. O segundo termo pode ser escrito como

$$\sum_{x_1} [\sum_{y_2} R_{n-1,\sigma}(T_f((s, X_1), Y_2).P(Y_2 = y_2)] P(X_1 = x_1)$$

$$= \sum V_{n-1,\sigma}(T_f((s, x_1)).P(X_1 = x_1) = E[V_{n-1,\sigma}(T_f(s, X_1))]$$

4.2.2. TEOREMA BÁSICO PARA O MODELO ESTOCÁSTICO I.I.D.

Para o modelo estocástico, o teorema básico, é enunciado como segue:

a) ***critério de optimalidade*** : Se f_n é um maximizador no estágio n para $1 \le n \le N$, então $(f_n)_N^1$ é óptima para o programa dinâmico de N períodos.

b) É valida a ***iteração de valor***

$$C_n(s) = \min\{r(s,a) + \beta E[C_{n-1}(T(s,a,X))]\}, n \in N, s \in S, a \in D(s)$$

O ganho obtido por período é muitas vezes, aleatório, isto é, depende de vários factores que possam ocorrer nesse período. Neste contexto, a função $r(s,a)$ será substituída por uma função finita $\tilde{r}(s,a,x)$, onde x é o valor da disturbância aleatória. Portanto, o nosso modelo de programação dinâmica alargar-se-à para um novo modelo que contém uma função finita, $\tilde{r}(s,a,x)$ sobre $D \times M$. Então a recompensa de N estágios para uma política π, quando o processo começa em s_0 é dada por

$$\tilde{R}_{N\pi}(s,Y) := \sum_{n=0}^{N-1} \beta^n \tilde{r}(\zeta_{n\pi}, \varphi_n(\zeta_{n\pi}), X_{n+1}) + \beta^N V_0(\zeta_N)$$

Assim $\tilde{V}_{n\pi}(s) := E[\tilde{R}_{N\pi}(s,Y)]$. Este modelo pode ser reduzido facilmente a um modelo como o anterior com a ajuda da seguinte proposição:

Proposição : Seja o modelo de programação dinâmica com $r(s,a) := E[\tilde{r}(s,a,X)]$

Então $\tilde{V}_{n\pi} = V_{N\pi}$ para todo o n e π. Como consequência, os dois modelos de programação estocásticos, terão a mesma função de valor, o mesmo conjunto de maximizadores na etapa n, $n \in N$, e o mesmo conjunto de políticas óptimas.

Exemplo

Problema:

Tem-se um recurso renovável, por exemplo peixe, que pode ser capturado em N períodos sucessivos. Assume-se que o recurso é medido numa determinada unidade e o produto de captura é vendido no mercado.

O problema é: qual deve ser o plano das capturas de modo a que no fim o lucro da venda seja máximo?

Formulação do modelo:

1. Espaço de Estado : S é o espaço de estados, s é a quantidade do recurso existente; claro que $s \geq 0$.

2. Espaço de Acções: A é o conjunto de acções a, sendo a a quantidade de recurso que fica depois da captura; claro que $a \leq s$.

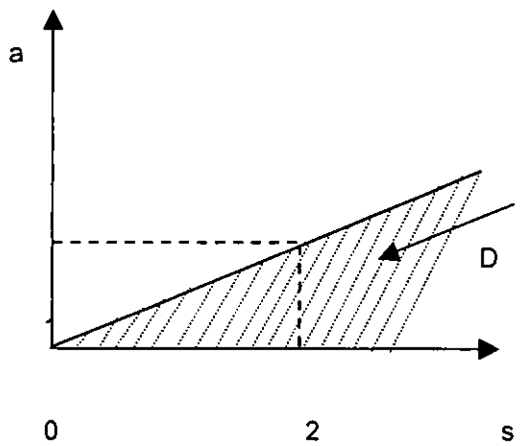

Fig. 2: Componentes do modelo

3. O conjunto de restrições é $D = (s,a)$: onde $a \leq s\}$

4. O conjunto de acções admissíveis no estado s, $D(s)$, é $\{a: a \leq s\}$. Quando $a = s$, significa que não se captura nada; quando $a = 0$, então foi capturado todo o recurso.

5. Função de Transição: A função de transição reflecte o crescimento do recurso que é função da quantidade a que ficou depois da captura e um conjunto de factores biológicos e ambientais governados estocásticamente. Podemos assumir que $T(s,a,X) = a.X$, sendo X uma variável aleatória.

6. A função de ganho é $r(s,a,X) = P.(s-a)$, onde P é o preço unitário do recurso e $(s-a)$ é a quantidade capturada.

7. A função de ganho terminal assumir-se-à identicamente igual a 0, isto é, $V_0(s) \equiv 0$.

8. O factor de desconto β é um número estritamente entre 0 e 1.

A iteração de valor será: $V_n(s) = \underset{a \in D(s)}{máx}\{ p(s-a) + \beta E V_{n-1}(T(s,a,X)) \}$, onde a esperança matemática é calculada de acordo com a distribuição da variável aleatória X.

4.3. ESTUDO DAS PROPRIEDADES ESTRUTURAIS DE UM MODELO DE PROGRAMAÇAO DINÂMICA

O estudo de propriedades estruturais de um modelo de programação dinâmica, consiste essencialmente em fazer uma análise das propriedades matemáticas das funções de valor $V_n(s)$ e das políticas óptimas tais como, a monotonia e concavidade [convexidade].

4.3.1. MONOTONIA DE FUNÇÕES DE VALOR $V_N(s)$

Poucos modelos de programação dinâmica possuem uma solução explícita. Portanto resultados qualitativos são de grande importância e, em particular, podem ajudar na simplificação dos cálculos. Neste âmbito, interessa-nos tratar questões do seguinte tipo:

1) Quando é que o ganho máximo no estágio n, $V_n(s)$, depende monotamente do horizonte n? Caso haja dependência, a existência de $V(s) := \lim_{n\to\infty} V_n(s)$ é assegurada e se $V(s)$ for finito, pode ser considerado como uma aproximação de $V_n(s)$ para um n muito grande.

2) Quando é que $V_n(s)$ depende monotamente do estado inicial s_0? Se este for o caso, as soluções numéricas (que sempre são disponíveis somente para um número finito de estados) podem ser achadas estudando a monotonia da função V_n.

4.3.1.1 DEPENDÊNCIA MONÓTONA DAS FUNÇÕES DE VALOR $V_N(s)$ EM RELAÇÃO AO HORIZONTE

O ponto de partida para investigação em programação dinâmica é a iteração de valor $V_n(s)$. Portanto, para o estudo da monotonia de $n \to V_n(s)$ vamos introduzir operadores apropriados que serão usados nas demonstrações dos teoremas.

Definição :

Os operadores L, U_f, para $f \in F$ e U transformam as funções v sobre S em funções Lv sobre D, $U_f v$ sobre S como se segue:

I. $(Lv)(s,a) := r(s,a) + \beta v(\mathrm{T}(s,a))$,

II. $(U_f)(s) := Lv(s,f(s)) := r_f(s) + \beta v(\mathrm{T}(s,a))$,

III. $(Uv)(s) :=$ sup $Lv =$ sup $U_f v(s)$, onde $a \in D(s)$ e $f \in F$.

Lema : Seja B um dos operadores L, U_f e U. Se $v \leq w$ então $Bv \leq Bw$.

Teorema: Se $V_0 \leq V_1$ [$V_0 \geq V_1$], então $n \longrightarrow V_n(s)$ é crescente [decrescente] para todo $s \in S$.

Demonstração: Vamos usar o método de indução matemática em n para fazer esta demonstração.

1) $V_0 \leq V_1$ segundo a condição do teorema;

2) Vamos supor que $V_k \leq V_{k+1}$, para $k \in N_0$; considerando que $k = n + 1$ teremos então $V_{n+1-1} \leq V_{n+1} \Rightarrow V_n \leq V_{n+1}$

4.3.1.2. DEPENDÊNCIA MONÓTONA DAS FUNÇÕES DE VALOR $V_N(s)$ EM RELAÇÃO AO ESTADO INICIAL S_0

Em muitos modelos de programação dinâmica com $S\subset(-\infty,+\infty)$ e $A\subset(-\infty,+\infty)$ as funções de valor e/ou alguns maximizadores são funções monótonas em relação ao estado inicial.

Às vezes a monotonia de $s \rightarrow V_n(s)$ é intuitivamente clara a partir da formulação do problema. Por exemplo em problema de investimento de capital, a função de valor cresce em relação ao capital inicial.

De seguida vamos abordar algumas propriedades de funções crescentes.

Lema : I. A composição de duas funções crescentes é uma função crescente.

II. A soma de duas funções crescentes, quando definidas, é crescente.

III. Se v é uma função crescente e $c \in R_+$, então $c.v$ é crescente.

IV. O produto de duas funções não negativas crescentes, é crescente.

A função v é crescente sse $1/v$ é decrescente.

Lema : (supremo de uma família de funções crescentes).

Seja S um conjunto ordenado e w uma função definida em D. Se a correspondência

$s \rightarrow D(s)$ é crescente e se $s \rightarrow w(s,a)$ é crescente para todo o a, então

$s \rightarrow w^*(s) := \sup\{ w(s,a) : a \in D(s)\}$ é crescente em S.

<u>*Demonstração*</u>:

Fixemos s e s' com $s \leq s'$. Se $a \in D(s)$, então também $a \in D(s')$, desde que $w(s', a) \geq w(s,a)$.

Assim,

$$w^*(s) = \sup \{ w(s,a) : a \in D(s) \} \leq \sup \{w(s',a) : a \in D(s) \} \leq \sup \{ w(s',a) : a' \in D(s')\} = w^*(s')$$

Teorema : Seja um S conjunto ordenado e vamos assumir que

a) A correspondência $D(.)$ é crescente, isto é, $D(s) \subset D(s')$ desde que $s \leq s'$;
b) A lei de transição T é crescente para todo a;
c) $R(.,a)$ é crescente para todo a;
d) V_0 é crescente;

Então $s \rightarrow V_n(s)$ é crescente para todo o $n \in N$

Demonstração: Vamos fazer esta demonstração usando a indução matemática.

$V_0(.)$ é crescente pela condição do teorema. Vamos assumir que $V_n(.)$ é crescente para todo $n \in N_0$. Então $V_n(\mathrm{T}(.,a))$ é crescente e por isso $LV_n(.,a)$ é também crescente pelo primeiro lema descrito na página 32. Como a correspondência $D(.)$ é crescente, podemos deduzir que V_{n+1} é crescente.

4.3.2. CONCAVIDADE E CONVEXIDADE DE FUNÇÕES DE VALOR V_N(s)

Uma das propriedades estruturais que é também importante estudar num modelo de programação dinâmica, é a concavidade [convexidade] de $s \rightarrow V_n(s)$ e/ou $s \rightarrow f_a(s)$. Antes de investigar esta propriedade, vamos fazer uma breve revisão de alguns conceitos que julgamos necessários para este estudo.

Definição:

Seja $v: I \rightarrow R$ definida num intervalo I, que é subconjunto do conjunto de números reais. Pode-se admitir que I seja um conjunto aberto, semi-aberto ou fechado, finito ou infinito:

a) A função $v: I \rightarrow R$ é côncava se $v[\lambda x + (1-\lambda)y] \geq \lambda v(x) + (1-\lambda)v(y)$ para todo x, $y \in I$ e $\lambda \in (0,1)$.

b) Se I é discreto, a função v é côncava se, se verificar a seguinte desigualdade:

$$2.v(i+1) \geq v(i) + v(i+2) \text{ para } i,\ i+1 \text{ e } i+2 \in I.$$

Nota: a função v é convexa se $-v$ é côncava.

Lema : (Propriedades das funções côncavas num intervalo I)

I. A soma de vária funções finitas côncavas é uma função côncava.

II. Se v é côncava e $c \in R_{+}$, então a função $c.v$ é côncava.

III. Se v é côncava então v é linearmente limitada, isto é existem constantes c e d tais que $v(i) \leq c + di$ para i $\in I$.

Se t é côncava em I, $t(i)$ é uma função contida num intervalo J (onde J é contínuo ou discreto se I é contínuo ou discreto), e se $v: J \longrightarrow R$ é concava [e crescente], então $v(t)$ é côncava [e crescente].

Lema : Seja w uma função côncava definida num conjunto convexo D. Então

$$s \longrightarrow w^{*}(s) := \sup w(s,.) \text{ é côncava, desde que } w^{*} < \infty.$$

<u>*Demonstração*</u>: Vamos escolher s, $s' \in S$ e $\alpha \in (0,1)$. Como D é convexo segundo a condição do lema, para $a \in D(s)$ e $a' \in D(s')$ teremos

$\alpha(s,a) + (1-\alpha).(s',a') = (\alpha s + (1-\alpha)s', \alpha a + (1-\alpha)a') \in D$. Daqui segue que pela concavidade de w

$w^{*}(\alpha s + (1-\alpha)s') \geq w(\alpha s + (1-\alpha)s', \alpha a + (1-\alpha)a') = w(\alpha.(s,a) + (1-\alpha).(s',a')) \geq \alpha. W(s,a) + (1-\alpha).w(s',a')$.

Como $a \in D(s)$ e $a' \in D(s')$, obtemos $w^{*}(\alpha s + (1-\alpha)s') \geq \alpha.w^{*}(s) + (1-\alpha).w^{*}(s')$.

Teorema : Vamos assumir que

I. D é convexo;

II. $D(s)$ é limitado para todo o s;

III. T é crescente para todo s;

IV. r e V_0 são côncavos;

Então V_n e W_n são finitos e côncavos para todo $n \in N$.

A demonstração é similar à que foi apresentada no lema anterior.

Quando se estuda a concavidade de V_n poderá também se estar interessado em saber se esta função é monótona ou não. Por isso, de seguida vamos enunciar um teorema que pode servir de ferramenta para a investigação das propriedades estruturais.

Teorema : (Concavidade de funções de valor crescentes):

Vamos supor que

i. D é convexo;

ii. $D(s)$ é limitado para todo o s;

iii. $D(.)$ é crescente;

iv. $T(.,a)$, $r(.,a)$ e V_0 são crescentes para todo o a,

v. T, r e V_0 são côncavos

Então V_n e $w_n(.,a)$ são finitos e côncavos (e V_n é crescente) para todo $n \in N$.

<u>*Demonstração*</u>:

As suposições (iii) e (iv) foram consideradas no teorema sobre a monotonia de

$s \longrightarrow V_n(s)$. Vamos verificar a concavidade de V_n usando método de indução em n.

Pela condição do teorema, para n=0, V_0 é concavo. Vamos supor que esta propriedade também se verifica para n-1 $\in N_0$. Então, por analogia, pode-se generalizar que $V_{n-1}(T(.,a))$ é côncavo em D devido a concavidade de T. A concavidade de w é preservada pela multiplicação V_{n-1} com uma constante β e adicionando a função concava em r. Existe para cada s uma função g em $D(s)$ tal que $w(s,.) \leq g$. A função $w(s,.)$ é limitada em $D(s)$, como $D(s)$ é limitado pela condição do teorema), por isso $V_n(s) = UV_{n-1}(s)$ é finito. Desta forma fica demonstrado que $V_n(s)$ é côncava pelo lema descrito na página 34.

4.3. MONOTONIA DOS MAXIMIZADORES

Nos modelos de programação dinâmica, dificilmente se consegue uma solução analítica e, para sair desta situação, procura-se obter a solução de programação dinâmica com ajuda do computador. Porém, por causa das limitações de memória, este método por vezes não é o mais viável. Daí que surge a necessidade de investigar propriedades das soluções de programação dinâmica que podem ajudar na simplificação do esforço computacional necessário na resolução dum problema de programação dinâmica. Por exemplo, quando o espaço de estados e o espaço das acções estão ordenados é útil saber se existe um maximizador, f_n, no estágio n que é crescente relativamente às duas relações em S e A. Se for este o caso, a maximização de $W_n(t,.)$ para $t \geq s$ pode ser restrita àquelas acções $a \in D(t)$ para as quais $a \geq f_n(s)$.

Nesta secção, serão apresentadas as condições sob as quais maximizadores apropriadamente escolhidos $f_a(s)$ são monótonos em s e/ou n.

Definição : uma correspondência $D(.)$ de S em A é chamada ***monotonamente completa*** se $s \leq s'$, $a \leq a'$ e $(s,a') \in D$, $(s',a) \in D$ implica que $(s,a) \in D$ e $(s',a') \in D$.

Lema : Seja $B(.)$ uma correspondência monotonamente completa de S em A. Se $f(s)$ for elemento menor ou maior de $B(s)$ para $s \in S$, então f é crescente.

Demonstração: Consideremos o caso em que $f(s)$ é o menor elemento de $B(s)$ pois, o outro caso é similar. Vamos supor que f não é crescente. Então existe $s' \geq s$ tal que $a := f(s) > a' := f(s')$. Mas $a' \in B(s)$ já que $B(.)$ é monotonamente completa e a é o menor elemento de $B(s)$, obtemos $a \leq a'$, que é uma contradição.

Se a função w em D tem um menor maximizador f, então f é crescente pelo lema descrito acima, sob condição de que a correspondência $s \rightarrow D^*(s)$:= conjunto de pontos máximos de $w(s,.)$ seja monotonamente completa. Este resultado é inútil a não ser que a última propriedade se verifique sem determinar os conjuntos $D^*(s)$. Felizmente a monotonia completa de $D^*(.)$, pode ser reduzida à monotonia completa da correspondência conhecida $D(.)$, desde que w tenha diferenças monótonas no sentido da seguinte definição.

Definição : Uma função finita w em D tem *diferenças crescentes* em s para todo o par de acções (a,a') com $a \leq a'$ se a função $s \rightarrow w(s,a) - w(s,a')$ é crescente em $D_a \cap D_{a'}$ sempre que o último conjunto seja não vazio.

Nota: A função finita w em D diz-se que tem diferenças crescentes se e só se

$$w(s,a) - w(s,a') \leq w(s',a) - w(s',a')$$

Um maximizador f é chamado o menor ou maior maximizador de w se $f(s)$ for o menor ou maior ponto máximo de $w(s,.)$ para todo s.

Teorema : sejam S e A conjuntos ordenados suponhamos que

I. $D(.)$ é monotonamente completo, isto é, se $s \leq s'$, $a \in D(s)$ e $a < a' \Rightarrow a' \in D(s)$ e $a \in D(s)$.

II. w tem diferenças crescentes, isto é, se $s < s'$ a função $s \rightarrow w(s,a) - w(s,a')$ é crescente.

Então cada maximizador maior ou menor de w é crescente.

Demonstração: Vamos mostrar que $B(.) := D^*(.)$ é monotonamente completo. Admitamos que $s \leq s'$, $a \in D^*(s)$, $a' \in D^*(s')$ e $a < a'$. Então também $a \in D(s)$ e $a' \in D(s')$. Sabe-se que $a' \in D^*(s)$ e $a \in D^*(s')$, como $D(.)$ é monotonamente completo. Considerando que a e a' são, respectivamente, os pontos máximos de $w(s,.)$ e $w(s',.)$, e como w tem diferenças crescente obtemos

$$0 \leq w(s,a) - w(s,a') \leq w(s',a) - w(s',a') \leq 0.$$

Isto implica que $a' \in D^*(s)$ e $a \in D^*(s')$, por isso $D^*(.)$ é monotonamente completo. Mas $D^*(.)$ é o conjunto de pontos máximos de $w(s,.)$, e , segundo lema descrito na página 36, f é crescente.

A função finita w tem diferenças decrescentes se $-w$ tem diferenças crescentes, isto é se $s \rightarrow w(s,a) - w(s,a')$ é decrescente em $D_a \cap D_{a'}$ sempre que $a < a'$. Assim f é minimizador de w sse f é maximizador de $-w$.

Corolário : Sejam S e A conjuntos ordenados e suponhamos que

I. $D(.)$ é monotonamente completo,

II. w tem diferenças decrescentes,

Então cada menor ou maior minimizador de w é crescente.

Demonstração: Sabe-se que $a' \in D(s)$ e $a \in D(s')$, pois, pela condição do corolário, $D(.)$ é monotonamente completo. Considerando que a e a' são, respectivamente, os pontos mínimos de $w(s,.)$ e $w(s',.)$, e como w tem diferenças decrescentes, isto é, $-w$ tem diferenças crescentes obtemos

$$0 \leq (-w(s,a)) - (-w(s,a')) \leq (-w(s',a)) - (-w(s',a')) \leq 0$$

$$0 \leq w(s',a') - w(s',a)) \leq w(s,a') - w(s,a) \leq 0$$

isto implica que $a' \in D^*(s)$ e $a \in D^*(s')$, por isso $D^*(.)$ é monotonamente completo. Mas $D^*(.)$ é o conjunto de pontos mínimos de $w(s,.)$, e segundo o lema descrito na página 36, f é crescente.

Se f_n é crescente, a solução do problema de maximização pode ser acelerada restringindo esta operação para valores de $a \geq f_n(s)$. Esta vantagem pode ser usada se f_n é uma função continua de Lipschitz com constante k, S e A discretos.

Denotaremos por ILIP(k) o conjunto de todas as funções f definidas em S que são crescentes continuas no sentido de Lipschitz em que k é a constante de Lipschitz, isto é,

$f(s) \leq f(t) \leq f(s) + (t-s).k$ para $s, t \in S$ e $s < t$.

Vamos de seguida mostrar como é que esta maximização se processa.

Consideremos que:

$|f_n(s') - f_n(s)| \leq k|s' - s|$ se $s' = s+1$, então $f_n(s') - f_n(s) > 0$

$f_n(s+1) - f_n(s) \leq k.1 \Rightarrow f_n(s+1) - f_n(s) \leq k$

a) para $k = 1 \Rightarrow f_n(s+1) \leq f_n(s) + 1$. é evidente que

$f_n(s) \leq f_n(s+1) \leq f_n(s) + 1$ ou $f_n(s+1) \in \{f_n(s), f_n(s) + 1\}$,

daqui podemos concluir que a pesquisa do ponto máximo no estado $s + 1$ fica apenas restringida para duas acções $a = f_n(s)$ e $a = f_n(s) + 1$.

b) Para $k = 2 \Rightarrow f_n(s+1) \leq f_n(s) + 2$ ou $f_n(s+1) \in \{f_n(s), f_n(s) + 1, f_n(s) + 2\}$

Neste caso, os possíveis pontos de máximos no estado $s + 1$ serão: $a = f_n(s)$, $a = f_n(s) + 1$ e $a = f_n(s) + 2$.

Portanto, analogamente, para qualquer k

$f_n(s+1) \in \{f_n(s), f_n(s) + 1, f_n(s) + 2, f_n(s) + 3, ..., f_n(s) + k + 1\}$ ou seja, a maximização no estado $s + 1$ terá lugar em um ou mais pontos:

$a = f_n(s)$, $a = f_n(s) + 1$, $a = f_n(s) + 2$, $a = f_n(s) + 3, ..., a = f_n(s) + k + 1$.

CAPÍTULO V

5. MODELO DE PROGRAMAÇÃO DINÃMICA ESTOCÁSTICA DE UM SISTEMA DE BÓNUS MALUS

O método de optimização sequencial de programação dinâmica é uma das ferramentas que possibilita a análise de um sistema de bónus malus. Neste trabalho é proposto o modelo de programação dinâmica estocástica, uma vez que o acidente pode ser considerado como acontecimento imprevisto ou fortuito do qual resulta um dano causado à coisa ou à pessoa e é muita das vezes aleatório.

5.1. FORMULAÇÃO DO PROBLEMA

Sejam assumidas as seguintes hipóteses:

1. Um condutor só pode ter no máximo um acidente por ano. Tendo em conta o crescente número de viaturas, bem como diferentes tipos de condutores, dos experientes aos sem experiência na condução, o que influência muito no tráfego, vamos considerar que a probabilidade de ter um acidente é 1/2 e a de não ter nenhum é também 1/2.

2. Um acidente tem custos aleatórios Y que dependem da gravidade do sinistro, e as probabilidades para que ocorra um acidente do custo K_i (i= 1,2,3) são:

$q(k_1) = 1/3$

$q(k_2) = 1/3$

$q(k_3) = 1/3$

Sendo $q(k_1) + q(k_2) + q(k_3) = 1$.

3. Assume-se que o condutor, perante um acidente, tem a liberdade de comunicar ou não o mesmo a Companhia Seguradora.
Se não comunicar (o que para a Companhia Seguradora significa que não teve acidente), passa para uma classe de prémio mais baixo em relação aquela em que se encontra actualmente; caso comunique, passa para uma classe com prémio mais alto.

4. Assuma-se a seguinte estrutura do sistema de bónus malus, em que teremos por um lado as diferentes classes em que o condutor pode-se encontrar e por outro lado os prémios correspondentes, denotaremos por i as classes e por $\pi(i)$ o respectivo prémio. De referir que a classe C_5 é a classe inicial no sistema, isto é, todo o condutor entra no sistema a partir desta classe.

Classe, (i)	Prémio de Classes, $\pi(i)$
C_1	$\pi(1) = 60$
C_2	$\pi(2) = 70$
C_3	$\pi(3) = 80$
C_4	$\pi(4) = 90$
C_5	$\pi(5) = 100$
C_6	$\pi(6) = 110$
C_7	$\pi(7) = 120$
C_8	$\pi(8) = 130$
C_9	$\pi(9) = 140$

Tabela *I*: *Estrutura do sistema de bónus malus*

Em seguida vamos apresentar, as regras de transição a que o condutor está sujeito de acordo com a acção tomada.

		Ano 1		Ano 2		Ano 3		Ano 4		Ano 5	
Classes	Prémios	0	1	0	1	0	1	0	1	0	1
C_9	140	C_8	C_9	C_7	C_9	C_6	C_9	C_5	C_9	C_4	C_9
C_8	130	C_7	C_9	C_6	C_9	C_5	C_9	C_4	C_9	C_3	C_9
C_7	120	C_6	C_8	C_5	C_9	C_4	C_9	C_3	C_9	C_2	C_9
C_6	110	C_5	C_7	C_4	C_8	C_3	C_9	C_2	C_9	C_1	C_9
C_5	100	C_4	C_6	C_3	C_7	C_2	C_8	C_1	C_9	C_1	C_9
C_4	90	C_3	C_5	C_2	C_6	C_1	C_7	C_1	C_8	C_1	C_9
C_3	80	C_2	C_4	C_1	C_5	C_1	C_6	C_1	C_7	C_1	C_8
C_2	70	C_1	C_3	C_1	C_4	C_1	C_5	C_1	C_6	C_1	C_7
C_1	60	C_1	C_2	C_1	C_3	C_1	C_4	C_1	C_5	C_1	C_6

Tabela *II* : *Quadro de transição de uma classe para outra.*

5. Assume-se que os prémios são anuais e que o condutor tem ligação com a companhia seguradora por cinco (5) anos, findos os quais esta cessa.

5.2. CONSTRUÇÃO DO MODELO

1. Espaço de estados

O espaço de estado S, representa as classes do sistema, isto é, S é o conjunto de todas as classes e têm o valor do prémio correspondente. Denota-se por $S = \{1,2,...5,...9\}$, onde 1 é a classe máxima de descontos, 9 é a classe máxima de agravamentos e 5 é a classe de entrada do condutor no sistema.

2. Espaço de acções

De acordo com a natureza do sistema de bónus malus, teremos como espaço de acções, o conjunto $A = \{0,1\}$, onde a acção 0, representa a situação em que o condutor não comunica à companhia seguradora algum acidente, a acção 1, representa a situação em que o mesmo informa sobre o acidente.

3. $D(s)$, pode ser considerado o conjunto de todas as acções admissíveis quando o condutor estiver numa determinada classe, $D(s) = \{0,1\}$.

4. Conjunto de Restrições

O conjunto de restrições é $D = S \times A$.

5. Função de Transição

A transição de uma classe para a outra, é feita por uma função $T(i,a)$, em que dependendo da acção que o condutor tomar na classe i, tanto pode transitar para a classe $i+1$ ou para a classe $i-1$.

$$T(i,a) = \begin{cases} i-1, \; Classe\, com\, menor\; prémio, & se\, a = 0 \\ i+1, \; Classe\, com\, maior\; prémio, & se\, a = 1; \end{cases}$$

6. Função de Custo

A função de custo é a base do nosso modelo, na medida em que é determinante para a escolha da acção óptima a tomar pois, por um lado exprime o custo de o condutor não informar à companhia seguradora preferindo pagar do seu próprio bolso e beneficiar dos descontos, por outro lado exprime o custo de informar à companhia seguradora e ter agravamentos no valor do prémio.

Este custo será calculado de acordo com as probabilidades de acontecer o acidente de custo k_l.

$$C_n(i) = \min_{a \in D(x)} \begin{cases} r(i,a) + Custo\,da\,Classe(i-1) + C_{n-1}(i-1)\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l), se\, a = 0 \\ Custo\,da\,Classe(i+1) + C_{n-1}(i+1)\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l), se\, a = 1 \;; \end{cases}$$

assume-se que $p(j)$ e $q(k_l)$ são independentes e $r(i,a) = \sum_{l=1}^{l=3}\sum_{j=0}^{j=1} k_l \times j \times p(j) \times q(k_l)$

Aqui, $p(j)$ = probabilidade de ocorrer j acidentes (j=0,1),

$q(k_l)$ = probabilidade de ocorrer um acidente de custo em k_l (l=1,2,3)

7. Factor de Desconto

É um número aleatório que se pode escolher entre 0 e 1 (estritamente) $0 < \beta < 1$

8. Função de Custo Terminal, C_0

A função do custo terminal, determina ou representa uma avaliação final no fim do horizonte de cinco (5) anos, quando cessa vínculo com a companhia seguradora, isto é, serve como instrumento para de uma forma geral, avaliar os ganhos ou recompensa máxima obtida durante a vigência do contrato, por isso,

$C_0(i) = 0$, para todo i=1, 2, ..., 9

5.3. OBTENÇÃO DA SOLUÇÃO

De acordo com o problema formulado e as premissas do modelo, vamos apresentar dois (2) casos que espelham as acções óptimas a serem tomadas pelo condutor, sendo para isso necessário, a minimização dos custos.

Para a minimização de custos no modelo estocástico, vamos usar o seguinte teorema básico:

a) ***critério de optimalidade*** : Se f_n é um maximizador no estágio n para $1 \leq n \leq N$, então $(f_n)_n^1$ é óptima para o programa dinâmico de N períodos.

b) É valida a ***iteração de valor***

$$C_n(s) = \min\{r(s,a) + \beta E[C_{n-1}(T(s,a,X))]\}, n \in N, s \in S, a \in D(s)$$

A iteração de valor será feita com base na comparação entre o custo de transitar para uma classe i-1 e o custo para transitar para uma classe i+1, donde teremos o cuidado de escolher sempre o valor mínimo que é obtido através da seguinte função de custo:

$$C_n(i) = \min_{a \in D(s)} \begin{cases} r(i,a) + Custo\,da\,Classe\,(i-1) + C_{n-1}(i-1)\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l), se\ a = 0; \\ Custo\,da\,Classe\,(i+1) + C_{n-1}(i+1)\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l), se\ a = 1\ ; \end{cases}$$

Note-se que, houve substituição de:

$r(i,a) = \sum_{l=1}^{l=3}\sum_{j=0}^{j=1} k_l \times j \times p(j) \times q(k_l)$, que representa o custo esperado de o condutor pagar por si próprio o acidente.

$E[C_{n-1}(T(i,a,X))]$ *por* $C_{n-1}(i-1)\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l)$, onde o $E = \sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l)$, representa a factor de aleatoriedade no sistema.

CASO 1

Vamos supor que os acidentes possíveis tenham os seguintes custos em unidades monetárias (*U.M.*):

$k_1 = 10$ *U.M.*

$k_2 = 20$ *U.M.*

$k_3 = 30$ *U.M.*

então, o custo médio que o condutor espera pagar do seu próprio bolso em caso de não anunciar o infortúnio será:

$$r(i,a)=\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} k_l \times j \times p(j) \times q(k_l) = 10U.M.$$

$$E=\sum_{l=1}^{l=3}\sum_{j=0}^{j=1} p(j) \times q(k_l) = 1$$ e, vamos considerar o factor de desconto $\beta=1$.

Substituindo na função de custo e comparando segundo a acção tomada pelo condutor teremos a seguinte tabela que nos indica os custos mínimos por ano.

	Classes (*i*)								
	1	2	3	4	5	6	7	8	9
$C_0(i)$	0	0	0	0	0	0	0	0	0
$C_1(i)$	70	70	80	90	100	110	120	130	140
$C_2(i)$	140	140	140	150	170	190	210	230	250
$C_3(i)$	210	210	210	210	220	230	250	280	310
$C_4(i)$	280	280	280	280	280	290	300	310	330
$C_5(i)$	350	350	350	350	350	350	360	370	380

Tabela *III*: *Custos mínimos ao Ano quando a classe de partida (no inicio do período de 5 anos) do condutor é i*

	Classes (*i*)								
	1	2	3	4	5	6	7	8	9
$C_1(i)$	0 ou 1	0	0	0	0	0	0	0	0 ou 1
$C_2(i)$	0	0	0	0	0	0	0	0	0
$C_3(i)$	0	0	0	0	0	0	0	0	0
$C_4(i)$	0	0	0	0	0	0	0	0	0
$C_5(i)$	0	0	0	0	0	0	0	0	0

Tabela *IV*: *Acções óptimas por ano*

CASO 2

Vamos analogamente, supor que os acidentes possíveis tenham os seguintes custos em unidades monetárias (*U.M.*):

$k_1 = 30$ *U.M.*

$k_2 = 50$ *U.M.*

$k_3 = 100$ *U.M.*

então, o custo médio que o condutor espera pagar do seu próprio bolso em caso de não anunciar o infortúnio será:

$$r(i,a) = \sum_{l=1}^{l=3} \sum_{j=0}^{j=1} k_l \times j \times p(j) \times q(k_l) = 30U.M.$$

$E = \sum_{l=1}^{l=3} \sum_{j=0}^{j=1} p(j) \times q(k_l) = 1$ e, vamos considerar o factor de desconto β=1.

Substituindo na função de custo e comparando segundo a acção tomada pelo condutor teremos a seguinte tabela que nos indica os custos mínimos por ano.

	Classes (*i*)								
	1	2	3	4	5	6	7	8	9
$C_0(i)$	0	0	0	0	0	0	0	0	0
$C_1(i)$	70	80	90	100	110	120	130	140	140
$C_2(i)$	160	160	160	180	200	220	240	260	280
$C_3(i)$	250	250	250	250	260	270	300	330	360
$C_4(i)$	340	340	340	340	340	350	360	370	390
$C_5(i)$	430	430	430	430	430	430	440	450	460

Tabela *V*: *Custos mínimos ao Ano quando a classe de partida (no inicio do período de 5 anos) do condutor é i*

	Classes (*i*)								
	1	2	3	4	5	6	7	8	9
$C_1(i)$	1	1	1	1	1	1	1	1	1
$C_2(i)$	0	0	0	0	0	0	0	0	0 ou 1
$C_3(i)$	0	0	0	0	0	0	0	0	0
$C_4(i)$	0	0	0	0	0	0	0	0	0
$C_5(i)$	0	0	0	0	0	0	0	0	0

Tabela *VI*: *Acções óptimas por ano*

CAPÍTULO VI

6. CONCLUSÕES E RECOMENDAÇÕES

Deste trabalho pode-se concluir que, para todos os casos, se os custos dos acidentes são mais baixos em comparação com os prémios, quase sempre a acção óptima do condutor é não anunciar o acidente, neste caso toma a acção, $a=0$, pois é preferível ele próprio suportar os custos do acidente a pagar prémios mais altos;

No caso 1, quando o condutor se encontra na classe 1 e 9, no último período do contrato, é lhe indiferente anunciar ou não o acidente, pois o prémio a pagar e os custos dos acidentes, são comparáveis. Analogamente o mesmo sucede no caso 2, quando se encontra na classe 9, no penúltimo período do contrato.

Verifica-se no caso 2, que no último período do contrato, o condutor prefere em qualquer classe i que esteja anunciar o acidente à companhia seguradora, portanto toma a acção, $a=1$, do que pagar do seu próprio bolso.

A análise dos respectivos casos, coloca o problema da necessidade da escala dos prémios ser ajustado aos custos do acidente, pois em todos os casos, a companhia seguradora vai receber menos em prémios, isto é, o condutor vai pagar cada vez menos em prémios, salvo quando o condutor se encontra no último período do contrato, no caso 2, que ele prefere remeter os acidentes à companhia seguradora, o que coloca o problema de se estabelecer a escala óptima de prémios.

Da análise às tabelas *III e V* de custos mínimos, vê-se que a função $C_n(i)$ dos custos mínimos esperados, é uma função crescente como função de i (n fixo) e de n (i fixo). Este facto corresponde à intuição: se o condutor se encontra numa classe de prémio mais alto no início do período de cinco (5) anos, os custos mínimos esperados são maiores do que se ele começar numa classe de prémio mais baixo. Analogamente, os custos mínimos esperados vão crescendo à medida que o tempo passa, para um mesmo estado ou classe i.

Seria interessante estudar:

a) O caso em que o condutor, pode causar mais do um (1) acidente por ano, devendo, diante de cada acidente, decidir se comunica ou não o mesmo à companhia seguradora. Esta modificação do modelo, aproximaria o mesmo à realidade;

b) A dependência do modelo em relação às probabilidades $p(j)$ e $q(k_l)$;

c) O problema de uma escala óptima de prémios no sistema de bónus malus, que não penaliza a companhia seguradora.

CAPÍTULO VII

7. BIBLIOGRAFIA

Bellman, R. (1957). Dynamic Programming, London.

Centeno, M.L.(1998). Teoria de risco. 4ª Edição. ISEG, UTL, Lisboa.

Cruz, B.L.M. (1997). Gestão de Recurso Renováveis. Trabalho de Licenciatura. UEM, Maputo.

Gnedenko, B.V. (1978). The Theory Of Probability. Mir Publisher, Moscow.

Guimarães, R.C., Cabral, J.A.S. (1997). Estatística. Editora McGraw-Hill de Portugal Lda.

Hinderer, K. (1991). Increasing Lipschitz Continous Maximizers of Some Dynamic Programs. Annals Of OR 29, Pág. 565 – 585.

Hinderer, K. (1993). Lecture Notes On Stochastic Dynamic Programming.

HTTP://www.certaseguros.com.br/dicionário. (17/10/03).

HTTP://www.geocities.com/jscmat/cadeias.htm (17/10/03).

HTTP://www.noroeste.com.br/história.htm (17/10/03).

HTTP://www.macseg.com.br/historia_do_seguro.htm (17/10/03).

Lemaire, J. (1985). Automobile Insurance: Actuarial Models, Kluwer Academic Publishers Group, Dordrecht.

Lemaire, J. (1995). Bónus malus Systems in Automobile Insurance, Kluwer Academic Publishers, Boston/ Dordrecht/ London.

Lemaire, J. (1998). "Bónus malus Systems: The European and Asian Approach to Merit-Rating". North American Actuarial Journal. **Volume 2, Number 1**, Pág. 26 - 47.

Triola, M. F. (1999). Introdução à Estatística. LTC-Livros Técnicos e Científicos, Editora S.A.

Ualane, N.R. (2003). Sistema de Bónus-Malus – Seguro Automóvel. Trabalho de Licenciatura. UEM, Maputo.