Programmation dynamique

Problème 1.

Une fable : Les héritiers

Il était une fois un riche industriel qui laissa à sa mort deux enfants à qui il légua à chacun 3 pièces d'Or et un testament où il était écrit:

Ce sera à chacun d'entre vous de décider chaque année ce qu'il fera de son argent; il pourra chaque année décider de ce qu'il investit jusqu'à la fin de la période et de ce qu'il conserve jusqu'à l'année suivante.

Le tableau suivant permettra de connaître à l'avance les rapports de ces investissements

	an 1	an2	an3
0 pièce	0	0	0
1pièce	4	3	2
2pièces	6	5	4
3pièces	9	6	5

A la fin de la troisième année, chacun fera le compte de qu'il possède, celui qui aura le plus aura droit à toute ma fortune.

Par exemple, celui qui investit la première année 1 pièce, puis la troisième 1 pièce aura à la fin

4+0+2+1(qui lui restait)=7 pièces.

3. Si j'arrive au début de la troisième année avec

debut 3eme annee $= x_3$	investissement $= u_3$	rapport $= g_3(x_3, u_3)$	fortune finale totale
0	0	0	0
1	0	0	1
	1	2	**2**
2	0	0	2
	1	2	$2+1=3$
	2	4	**4**
3	0	0	3
	1	2	$2+2=4$
	2	4	$\mathbf{4+1=5}$
	3	5	**5**

Donc dans chacun des cas voici la meilleure décision et la fortune totale prévue

debut 3 eme annee $= x_3$	bon choix $= u_3^*$	fortune maximale possible $= J_3(x_3)$
0	0	0
1	1	2
2	2	4
3	2 ou 3	5

2. Si j'arrive au début de la deuxième année avec

debut 2eme annee $= x_2$	investissement $= u_2$	il reste x_3
0	0	0
1	0	1
	1	0
2	0	2
	1	1
	2	0
3	0	3
	1	2
	2	1
	3	0

Panorama année 2 +année 3:

$g_2(x_2,u_2)$ ou plus simplement $g_2(u_2)$ représentera le rapport qu'apporte le choix d'investir u_2 au cours de la période 2;

bien sûr, si on commence la période 2 avec x_2, si on investit u_2 il restera $x_3=x_2-u_2$ pour la période 3.

debut 2eme annee $= x_2$	investissement $=u_2$	rapport $= g_2(u_2)$	x_3	$J_3(x_3)$	$g_2(u_2) + J_3(x_3)$
0	0	0	0	0	$\mathbf{0}$
1	0	0	1	2	$0 + 2$
	1	3	0	0	$\mathbf{3+0}$
2	0	0	2	4	$0 + 4$
	1	3	1	2	$\mathbf{3+2}$
	2	5	0	0	$\mathbf{5+0}$
3	0	0	3	5	$0 + 5$
	1	3	2	4	$\mathbf{3+4}$
	2	5	1	2	$\mathbf{5+2}$
	3	6	0	0	$6 + 0$

Concluons pour l'année 2

debut 2e annee $= x_2$	invest $=u_2$	rapport $= g_2(u_2)$	x_3	$J_3(x_3)$	$g_2(u_2) + J_3(x_3)$	u_2^*	$J_2(x_2)$
0	0	0	0	0	$\mathbf{0}$	0	0
1	0	0	1	2	$0 + 2$		
	1	3	0	0	$\mathbf{3+0}$	1	3
2	0	0	2	4	$0 + 4$		
	1	3	1	2	$\mathbf{3+2}$	1 ou2	5
	2	5	0	0	$\mathbf{5+0}$	1ou2	5
3	0	0	3	5	$0 + 5$		
	1	3	2	4	$\mathbf{3+4}$	1ou2	7
	2	5	1	2	$\mathbf{5+2}$	1ou2	7
	3	6	0	0	$6 + 0$		

1. Au début de la première année

Nécessairement x_1=3

debut 1 annee $= x_1$	invest $= u_1$	il reste x_2
3	0	3
	1	2
	2	1
	3	0

debut 1 annee $=x_1$	invest $=u_1$	$g_1(u_1)$	il reste x_2	$J_2(x_2)$	$g_1(u_1)+J_2(x_2)$	u_1^*	$J_1(3)$
3	0	0	3	7	$0+7$		
	1	4	2	5	$4+5$		
	2	6	1	3	$6+3$		
	3	9	0	0	$9+0$	$1,2,3$	9

Conclusion:

Le meilleur total sera 9, obtenu de diverses manières.

On dira que

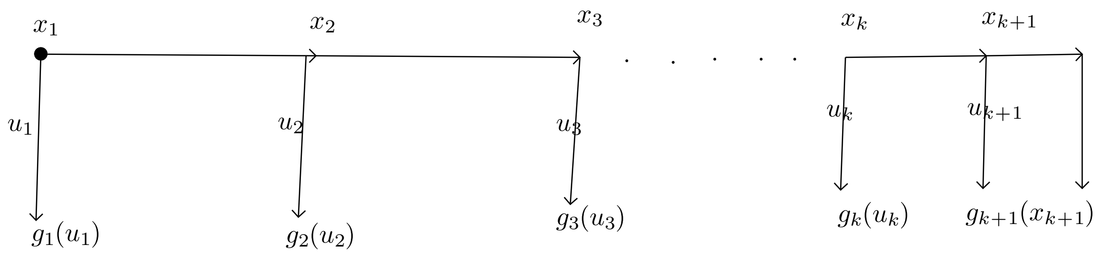
le gain maximal est $J_1(3)=9$

la politique optimale est $u_1^* = , u_2^* = , u_3^* = .$

1. Objet et principes

On considère un système dynamique à temps discret et une fonction de gain additive dans le temps; le système décrit l'évolution de l'état du processus au fil de n périodes et a la forme $\mathrm{x}_{n+1}=f_n\big(x_n, u_n\big)$ où x_n est l'état du système au début de la période n, u_n la décision (ou le contrôle) prise à la période n.

La fonction de gain est associée à chaque période du système $g_n(x_n, u_n)$ et ces gains s'additionnent de telle sorte qu'à l'issue du processus leur somme est le coût total $g_{n+1}(x_{n+1})+\sum_{k=1}^{n} g_k(x_k, u_k)$.



La fonction de gain est associée à chaque période du système $g_n(x_n, u_n)$ et ces gains s'additionnent de telle sorte qu'à l'issue du processus leur somme est le coût total $g_{n+1}(x_{n+1}) + \sum_{k=1}^{n} g_k(x_k, u_k)$

L'outil :

Théorème 1. *Le principe d'optimalité de Bellman:*

Soit un problème de décisions séquentielles sur n périodes, S_{n+1} *l'ensemble des états possibles; alors pour tout état initial* x_1 *le coût minimal* $J^*(x_1)$ *est égal à* $J_1(x_1)$*, où la fonction* J_1 *est définie par l'algorithme suivant:*

$J_{n+1}(x_{n+1}) := g_{n+1}(x_{n+1})\ \forall x_{n+1} \in S_{n+1}$

et pour tout $k = n, n-1, \ldots, 2, 1$ $\quad J_k(x_k) := \max\left\{ g_k\big(x_k, u_k\big) + J_{k+1}(f_k(x_k, u_k)), u_k \,\text{possibles}\right\}$.

C'est à dire: soit la « politique » $\pi^* = (u_1^*, \ldots, u_n^*)$, si celle-ci est optimale alors pour tout i la politique $(u_i^*,\ \ldots,\ u_n^*)$ est optimale pour le sous-problème débutant dans l'état x_i.

Remarque 2.

Identifier les états, les étapes et les règles régissant le passage d'une étape à l'autre (et donc les changements d'états en fonction des décisions prises).

2. Exemples

2.1 Le sac à dos (hyper classique)

On considère un assortiment comportant N types d'objets, numérotés de 1 à N, de valeurs respectives $(c_1, \ldots, c_N)$, de volumes respectifs $(a_1, \ldots, a_N)$ pour le transport desquels on dispose d'un « sac à dos » de volume b; on cherche donc à définir les quantités respectives $(u_1, \ldots, u_N)$ de ces divers types de telle sorte que:

i) $\sum_{k=1}^{N} c_k u_k$ soit maximal

ii) $(u_1, \ldots, u_N) \in \mathbb{N}^N$

iii) $\sum_{k=1}^{N} a_k u_k \leqslant b$.

On peut modéliser ce problème comme un processus de décisions en N étapes.

Pour chaque k

* l'état x_k correspond à l'espace disponible pour les objets de type k,k+1,..,N

* la décision u_k consistant à choisir le nombre d'objets de type k

* la fonction de transition fonctionne comme suit:

si u_k objets de type k sont sélectionnés, l'espace disponible pour l'étape k+1 est $x_{k+1}=f_k\big(x_k, u_k\big)=x_k-a_k u_k$.

et le profit total est $\sum_{k=1}^{N} g_k\big(x_k,\ u_k\big) = \sum_{k=1}^{N} c_k u_k$, que l'on doit maximiser par rapport à $(u_1, \ldots, u_N)$.

Résolution:

On commence par $J_N(x_N)$

c'est le maximum des valeurs de $g_n(x_n, u_n) = g_n(u_n) = c_N u_N$;

u_N c'est la quantité d'objets N emportés , donc il faut que $x_N - a_N u_N \geqslant 0$.

$J_N(x_N) = \max\{c_N u_N, x_N - a_N u_N \geqslant 0, u_N \in \mathbb{N}\}$, il est clair que la décision optimale est $u_N^* = E\left(\frac{x_N}{a_N}\right)$.

Puis (on « remonte » à la catégorie N-1):

On veut trouver $J_{N-1}(x_{N-1})$

$J_{N-1}(x_{N-1}) = \max\left\{c_{N-1}u_{N-1} + J_N(x_{N-1} - a_{N-1}u_{N-1}), \text{tels que } u_{N-1} \leqslant \right.$
$\left. \leqslant \frac{x_{N-1}}{a_{N-1}} \text{ avec } u_{N-1} \in \mathbb{N}\right\}$

(car u_{N-1} doit être un entier positif tel que $x_N = x_{N-1} - a_{N-1}u_{N-1}$ soit positif).

Donc pour chaque valeur de x_{N-1} (entre 0 et b)

il faut considérer chaque valeur de u_{N-1} (inférieure ou égale à $E\left(\frac{x_{N-1}}{a_{N-1}}\right)$)

et pour chacune $x_N = x_{N-1} - a_{N-1}u_{N-1}$ et $u_N^* = E\left(\frac{x_N}{a_N}\right)$.

Ainsi on déduit pour chaque valeur de x_{N-1} la meilleure décision u_{N-1}^* et $J_{N-1}(x_{N-1})$.

etc....

jusqu'à x_1 et $J_1(x_1)$:

Comme x_1 est le premier état il vaut nécessairement b et $J_1(x_1) = J_1(b) = \max\left\{c_1u_1 + J_2(x_1 - a_1u_1), x_1 - a_1u_1 \geqslant 0, u_1 \in \mathbb{N}\right\}$; nous opérerons comme plus haut pour chaque valeur de u_1 (inférieure ou égale à $E\left(\frac{x_1}{a_1}\right)$, ceci déterminera la meilleure décision u_1^* et $J_1(x_1)$.

En récapitulant les résultats précédents nous obtiendrons alors (u_1^*, ..., u_N^*).

Tout cela exige des tableaux

Traitons un exemple numérique

N=3, b=6,

objet	1	2	3
valeur	5	3	7
volume	3	2	4

- L'étude de u_3 donne

x_3	0	1	2	3	4	5	6
u_3	0	0	0	0	$0,1$	$0,1$	$0,1$
$J_3(x_3)$	0	0	0	0	$0,7$	$0,7$	$0,7$
$u_3^*(x_3)$	0	0	0	0	1	1	1

et alors on peut déterminer $J_3(x_3)$

x_3	0	1	2	3	4	5	6
$J_3(x_3)$	0	0	0	0	7	7	7
$u_3^*(x_3)$	0	0	0	0	1	1	1

-

Passons à J_2: $J_2(x_2) = \max \left\{ 3u_2 + J_3(x_2 - 2u_2), x_2 - 2u_2 \geqslant 0, u_2 \in \mathbb{N} \right\}$

D'où le tableau de la deuxième étape:

x_2	0	1	2	3	4	5	6
u_2	0	0	$0;1$	$0;1$	$0;1;2$	$0;1;2$	$0;1;2;3$
$J_2(x_2)$							

; d'où

$J_2(0) = \max \{3 \times 0 + J_3(0)\}$=0

$J_2(1) = \max\{3 \times 0 + J_3(1)\}$=0

$J_2(2) = \max\{3 \times 0 + J_3(2), 3 \times 1 + J_3(0)\}$=$\max\{0 + 0, 3 + 0\} = 3$ et $u_2^*(2) = 1$

$J_2(3) = \max\{3 \times 0 + J_3(3), 3 \times 1 + J_3(1)\}$=$\max\{0 + 0, 3 + 0\} = 3$ et $u_2^*(3) = 1$

$J_2(4) = \max\{3 \times 0 + J_3(4), 3 \times 1 + J_3(2), 3 \times 2 + J_3(0)\}$=$\max\{0 + 7, 3 + 0, 6 + 0\} = 7$ et $u_2^*(4) = 0$.

$J_2(5) = \max\{3 \times 0 + J_3(5), 3 \times 1 + J_3(3), 3 \times 2 + J_3(1)\}$=$\max\{0 + 7, 3 + 0, 6 + 0\} = 7$ et $u_2^*(5) = 0$

$J_2(6) = \max\{3 \times 0 + J_3(6), 3 \times 1 + J_3(4), 3 \times 2 + J_3(2), 3 \times 3 + J_3(0)\}$= $\max\{0 + 7, 3 + 7, 6 + 0, 9 + 0\} = 10$ et $u_2^*(6) = 1$.

D'où le tableau récapitulatif

x_2	0	1	2	3	4	5	6
$J_2(x_2)$	0	0	3	3	7	7	10 .
$u_2^*(x_2)$	0	0	1	1	0	0	1

Passons à J_1:

Par définition $x_1 = 6$ et donc $J_1(6) = \max\left\{5u_1 + J_2(6 - 3u_1), 6 - 3u_1 \geqslant 0, u_2 \in \mathbb{N}\right\}$, c'est à dire

$J_1(6) = \max\{5 \times 0 + J_2(6), 5 \times 1 + J_2(3), 5 \times 2 + J_2(0)\}$=$\max\{0 + 10, 5 + 3, 10 + 0\} = 10$.

D'où $u_1^*(x_1)$=0 ou 2.

Si on prend $u_1^*(6)$=0 alors $x_2 = 6$ il faut choisir $u_2^*(6)$=1 et $x_3 = 4$ et il faut choisir $u_3^*(4)$=1 ; la « politique » optimale de décisions est donc (0,1,1) et le gain maximal est 10.

Si on prend $u_1^*(6)$=2 alors de même on déterminera la « politique » optimale de décisions est donc (2,0,0) et le gain maximal est 10 aussi.

Problème 2. (toute résolution par une démarche autre que la programmation dynamique sera refusée)

Un randonneur doit préparer son sac à dos; il ne peut emporter plus de 9 Kgs et dispose de deux types d'éléments, de valeur nutritive variable, conditionnés par unités non fractionnables.

Aliments	type 1	type 2
poids unitaire(kgs)	7	5
quantités disponibles	4	3
valeur nutritive	15	10

Déterminer ce qu'il doit emporter dans son sac à dos afin d'obtenir la meilleure valeur nutritive.

3. Travaux dirigés

Exercice 1. Ce touriste est aussi un commerçant

Ah que les produits high-tech sont bon marché dans cette boutique duty-free, qu'il serait rentable d'en acheter afin de les revendre au retour à Paris; je peux transporter jusqu'à 10 kg (car je voyage en charter low cost) et voici les infos

item	poids	benefice attendu
1	2	110
2	3	160
3	5	260
4	4	210

.

Qu'allez-vous emporter afin d'obtenir le meilleur gain?

Exercice 2. Conserver ou remplacer le matériel

Une usine chimique utilise un équipement extrémement sensible à la corrosion, ce qui diminue sa productivité.

Lorsque l'équipement a une ancienneté de t années (x=0,..,4) le revenu net (après déduction des frais) est égal à $I(x) = 26 - 2x - \frac{x^2}{2}$; si l'ancienneté est x> 4 alors I(x)=0; le coût de remplacement est égal à 22 et le matériel, mis au rebut, n'est pas revendable.

On désigne par 0 l'année d' installation de l' équipement (le 1 septembre) et la décision de le conserver ou de le remplacer se prend chaque année à la rentrée, c'est à dire début septembre; on désignera pour chaque année (i=1,..,5) par u_i=0 la décision (à l'issue de la i-ème année) de conserver l'équipement et par u_i=1 la décision de renouveler l'équipement, le revenu net sera donc $\sum_{i=1,\ldots,5} g(x_i, u_i)$, où $g(x_i, u_i) = 26 - 2x_i - \frac{x_i^2}{2}$ ou $g(x_i, u_i)$=26-22=4, suivant la valeur de u_i et l'ancienneté de l'équipement x_i.

Quelle la politique la plus rentable pour les 5 années à venir?

1. Exprimer, suivant la valeur de u_5 et de x_5, la valeur de $g(x_5, u_5)$ et , si on désigne par $J_i(x_i)$ le profit maximal lorsque l'équipement a une ancienneté égale à x_i à l'issue de la 5 ième année, déterminer $J_5(x_5)$ pour x_5=1,2,3,4,5.
2. On posera $x_{i+1}=f(x_i, u_i)$, exprimer la fonction f.
3. Comme $J_4(x_4) = \max\{g(x_4, u_4) + J_5(x_4), u_4 = 0, 1\}$, déterminer $J_4(x_4)$ pour x_4=1,2,3,4
4. De même calculer $J_3(x_3)$ pour x_3=1,2,3
5. Et $J_2(x_2)$ pour x_2=1,2
6. Et $J_1(x_1)$ pour x_1=1
7. Déterminer la politique à suivre pour ces 5 années avoir d'obtenir un gain maximal.

Exercice 3.

Une entreprise dispose d'un budget de 500.000 Euros pour l'amélioration de ses 3 sites industriels durant l'année à venir; les sommes investies seront chaque fois un multiple entier de 100.000 Euros; les retours espérés en fonction de la somme investie sont donnés en multiples entiers de 100.000 Euros dans le tableau ci-joint

Site $\downarrow$et investissement $\longrightarrow$	0	1	2	3	4	5
1	1	3	4	5	5.5	6
2	0.5	2.5	4	5	6	6.5
3	2	4	5.5	6.5	7	7.5

Le problème comportera 3 étapes correspondant à la décision d'investissement pour les 3 sites, 1,2,3; à l'étape k la variable d'état x_k représente (tout comme dans le cas du sac à dos) le montant disponible pour investissement dans les sites k,k+1,.,3; l'évolution de l'état du système sera donnée par le formule $\forall k \in \{1,\ 2,\ 3\}$, $x_{k+1} = f_k(x_k,\ u_k) = x_k$-$u_k$, où représente l'investissement décidé pour le site k.

Chaque état x_k appartient à $S_k = \{0,\ 1,\ \ldots,\ 5\}$ et pour chaque état les décisions *admissibles* u_k appartiennent à $\{0,\ 1,\ \ldots,\ x_k\}$; si on note $g_k\big(x_k, u_k\big)$ le bénéfice anuel obtenu en investissant sur le site k la somme u_k (en fait dans cet exemple $g_k\big(x_k,\ u_k\big)$ ne dépend que de u_k, mais attention: il faut $x_k - u_k \geqslant 0$) alors le bénéfice total sera

$g_4(x_4) + \sum_{k=1}^{3} g_k(x_k, u_k)$ en notant x_4 la partie des 500.000 Euros non investie.

On peut supposer que $g_4(x_4){=}x_4$ (la somme immobilisée sur notre compte en banque ne rapporte rien); nous recherchons donc à optimiser $x_4 + \sum_{k=1}^{3} g_k(x_k, u_k)$.

Problème 3. La compagnie Iciçaenjette est catégorique: au maximum 35 Kg de bagages.

Vous voudriez ramener de votre séjour à Las vegas

1) des météorites du nevada: poids unitaire 800 g, valeur 155 Euros

2) des équipements de cow-boy: poids unitaire 1200 g, valeur 210 Euros

3) des souvenirs indiens : poids unitaire 950 g, valeur 180 Euros

Remarque 3. Complexité ?

- Nombre d'étapes :n+1

- Pour chaque étape : le nombre d'états, le nombre de calculs induits par chaque état.

Objectifs:

1. Savoir reconnaître un système dynamique à temps discret et une fonction de coût ou de gain additive dans le temps.

2. Savoir déterminer les états et identifier la fonction qui définit l'état suivant en fonction de l'état actuel et de la décision prise à l'instant actuel.

3. Savoir déterminer l'expression de la fonction de coût ou de gain à chaque étape.

4. Savoir calculer pour chaque étape le coût (ou le gain) total en fonction de l'état et de la décision prise à cette étape en cet état.

5. Savoir faire la synthèse et déterminer la politique optimale (cad la suite des bonnes décisions).