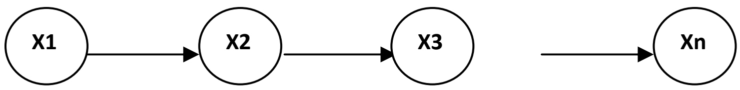
Chapitre 1

Programmation Dynamique Stochastique

Exemple introductif

Un système **dynamique** est un système qui évolue au cours du temps ; nous nous intéressons aux cas d'une évolution **discontinue** dans le temps qu'on peut schématiser comme suit :



Où : x_k désigne l'état du système au début de la période k

Si on considère un problème de gestion de stocks

- L'état du système : correspond au nombre d'unités disponibles d'un certain article.

$X_{k+1}=X_k$ + nombre_unités_commandées – **nombre_unités_demandées[1].**

Dans la réalité:

Question 1 : Etant donné l'état du système dans la période k ; peut-on savoir **avec certitude** quel sera l'état du système au début de la période k+1 ? Pourquoi ?

- Si on considère les coûts suivants :
  - C1 : Coût de stockage associé aux unités en excès en fin d'une période
  - C2 : Coût de pénurie encouru pour les demandes n'ayant pu être satisfaites pendant une période.

Question 2 : Peut-on savoir, au début d'une période, **avec certitude** quel coût (C1 ou C2) sera marqué en fin de cette période? Pourquoi ?

Dans tout çà, on est devant un système dynamique dont l'évolution est stochastique du moment que les demandes sont des variables aléatoires.

---

[1] Les demandes sont des variables aléatoires indépendantes suivant des lois de probabilités données.

1. Introduction

Pour les systèmes déterministes, les règles de conduite optimales et les décisions associées sont obtenues en minimisant les coûts totaux. Les méthodes de programmations dynamiques basées sur des modèles déterministes supposent que le coût et le changement de l'état résultants de chaque décision, même pour leurs valeurs futures, sont connus avec certitude. Malgré la simplicité attirante que possède ces méthodes, ceci n'est le cas ni dans nos vies courantes ni dans le monde des affaires ni dans le domaine scientifique ou de l'ingénierie.
Dans ce cours, nous introduirons une approximation plus proche de la réalité et nous supposerons que chaque décision peut produire un certain nombre de réalisations possibles. Nous supposerons que chacune de ces réalisations possède une probabilité connue a priori. Pour un modèle plus réel, ces probabilités ne sont pas connues a priori mais peuvent être estimées.

2. Le Modèle de base

Les problèmes auxquels nous nous intéresserons sont associés à des situations où des décisions doivent être prises de manière séquentielle, les conséquences de chaque décision n'étant pas toujours parfaitement maîtrisées mais pouvant être anticipées jusqu'à un certain point avant que la prochaine décision ne soit prise. Le but recherché est la minimisation d'un coût (ou la maximisation d'un profit) associé à la suite de décisions retenues et à leurs conséquences.

Le modèle de base, considéré dans ce contexte, repose sur deux éléments essentiels :

1- Un système dynamique à temps discret.
2- Une fonction coût additive dans le temps.

Le système dynamique décrit l'évolution du processus au fil de N période et a la forme :

$$x_{k+1} = f_k(x_k, u_k, \omega_k), \qquad k = 1,2,\dots,N$$

Où,

k : numérote les périodes (ou les étapes) du processus.

f_k : appelés fonctions de transfert ou de transition.

x_k : état du système au début de la période k , $x_k \in S_k$.

u_k : décision (ou contrôle) devant être prise à la période k, $u_k \in C_k$; une variable de décision est généralement contrainte à prendre ses valeurs dans un sous ensemble non vide $U_k \subseteq C_k$ qui peut dépendre de x_k.

ω_k : paramètre aléatoire, parfois appelé bruit ou perturbation ; $\omega_k \in D_k$ et est caractérisé par une distribution de probabilité pouvant dépendre explicitement de x_k et de u_k et non pas des perturbations précédentes.

μ_k : associe à chaque état $x_k \in S_k$ une décision $u_k = \mu_k(x)$.

Une fonction coût $g_k(x_k, u_k, \omega_k)$ est associée à chaque période du système dynamique. En outre, un coût terminal dépendant de l'état atteint à la fin du processus est défini et noté par $g_{N+1}(x_{N+1})$

π : politique de décision qui est une suite de fonctions $\pi = \{\mu_{1,}\mu_{2,\dots,}\mu_N\}$

L'espérance du coût total associé à x_j *et* à π est définie comme suit :

$$J_\pi(x_j) = E[g_{N+1}(x_{N+1}) + \sum_{k=j}^{N} g_k(x_k, \mu_k(x_k), \omega_k)]$$

π^* : dénote la politique optimale ; π^* est optimale pour un état initial x_1 si elle vérifie :

$$J_{\pi^*}(x_1) = \min_{\pi \in \Pi} J_\pi(x_1)$$

Le coût optimal pour un système débutant dans l'état x_1, noté $J^*(x_1)$, satisfait :

$$J^*(x_1) = J_{\pi^*}(x_1) = \min_{\pi \in \Pi} J_\pi(x_1)$$

Note :

Les processus de décisions déterministes ne font pas intervenir de paramètres aléatoires. Dans le cadre du modèle de base présenté ci-dessus, ils correspondent à des systèmes où chaque perturbation ω_k est une variable dégénérée, c'est-à-dire ne prenant qu'une seule valeur.

Exemple1 : Modélisation du problème du sac à dos en tant qu'un processus de décisions séquentielles déterministes.

Considérons un assortiment comportant N types d'objets numérotés de 1 à N, chaque objet de type k étant caractérisé par une valeur $c_k \geq 0$ entière et un volume $a_k > 0$ également entier. Le problème de la sélection, parmi cet assortiment, d'un sous ensemble d'objets de valeur maximale ne dépassant pas un volume globale ***b*** donné peut être modélisé par un processus de décisions caractérisé par :

Etapes

Numérotés de 1 à N (*ici étape correspond à objet*).

<u>Décisions</u>

La décision u_k de l'étape k consiste à choisir le nombre d'objets de type k à inclure dans la sélection.

Chaque décision doit vérifier : $0 \leq u_k \leq \left\lfloor \frac{x_k}{a_k} \right\rfloor$ entier.

<u>Etats</u>

L'état x_k du système au début de l'étape k correspond à l'espace réservé pour les objets de type : *k*, *k+1*,…,*N*. $x_k \in \{0,1, \dots, b\}$.

<u>Fonction de transfert (transition)</u>

Si u_k objets de type k sont sélectionnés, l'espace disponible à l'étape k+1 est :

$$x_{k+1} = f_k(x_k, u_k) = x_k - a_k u_k, \quad k=1,2,\dots,N$$

<u>Fonction coût</u>

La valeur des objets sélectionnés à l'étape k est $g_k(x_k, u_k) = c_k u_k$, et le profit totale est $\sum_{k=1}^{N} g_k(x_k, u_k) = \sum_{k=1}^{N} c_k u_k$

3. La Programmation dynamique

3.1 Définition et principe

La programmation dynamique est une technique algorithmique qui permet de résoudre une catégorie particulière des problèmes d'optimisation nécessitant des séquences de décisions corrélées. Elle a été désignée par ce terme pour la première fois dans les années 1940 par Richard Bellman.

La programmation dynamique repose sur une observation très simple, connue sous le nom de principe d'optimalité dont la formulation est la suivante :

Théorème

Soit $\pi^* = \{\mu_1^*, \mu_2^*, \dots, \mu_N^*\}$ une politique optimale pour un problème de décisions sur N périodes. Si en mettant en œuvre π^*, l'état x_j est atteint à la période j, la politique partielle $\{\mu_j^*, \mu_{j+1}^*, \dots, \mu_N^*\}$ est optimale pour le sous-problème sur (N-j+1) périodes, *débutant à l'étape j dans l'état x_j et se terminant au début de l'étape N+1. En d'autres termes, la politique partielle $\{\mu_j^*, \mu_{j+1}^*, \ldots, \mu_N^*\}$ minimise l'espérance des coûts :*

$$E[g_{N+1}(x_{N+1}) + \sum_{k=j}^{N} g_k(x_k, \mu_k(x_k), \omega_k)]$$

3.2 Algorithme de la programmation dynamique déterministe

1. *Construction de la politique optimale*
   a. *Initialisation*

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

   b. *Pour k=N, N-1,…,1 et pour tout $x_k \in\ S_k$ résoudre*

$$J_k\,(x_k) := \min_{u_k \in U_k(x_k)} \{\, g_k(x_k, u_k) + J_{k+1}\,(f_k(x_k, u_k))\, \}$$

   Et stocker la valeur $J_k\,(x_k)$ ainsi que la décision $\mu_k^*(x_k) \coloneqq u_k^*$ pour laquelle le minimum est atteint.

2. *Lecture de la solution optimale (pour l'état initial x_1)*
   a. *Le coût optimal est $J^* = J_1(x_1)$.*
   b. *La suite de décisions optimales est $u_1^* = \mu_1^*(x_1)$*

   Et $u_k^* = \mu_k^*(f_{k-1}(x_{k-1}, u_{k-1}^*)),\ k = 2,3, \ldots, N$

Exemple Résolution du problème de sac à dos par l'algorithme de la PD

On cherche le chargement optimal d'un sac de volume 6. Les caractéristiques des objets susceptibles d'être inclus dans ce sac sont données par le tableau suivant :

Objet K	1	2	3
Valeur c_k	5	3	7
Volume a_k	3	2	4

A chaque étape ***k*** nous cherchons la solution du problème suivant :

$$J_k\,(x_k) := \max_{u_k \in U_k(x_k)} \{\, c_k\, u_k + J_{k+1}\,(x_{k+1})\, \}$$

S.C $x_{k+1} = x_k - a_k u_k \geq 0$

$$u_k \geq 0\ entier$$

Etape3

$J_3\,(x_3) := \max\limits_{u_3}\{\, c_3\, u_3 \,\} = \max\limits_{u_3}\{7u_3\}$

S.C $0 \le u_3 \le \left\lfloor \frac{x_3}{4} \right\rfloor$ *entier*

* Pour $x_3 \in \{0{,}1{,}2{,}3\}$: $J_3\,(x_3) = \max_{u_3=0}\{7u_3\} = 0$; $\mu_3^*(x_3) = 0$

* Pour $x_3 \in \{4{,}5{,}6\}$: $J_3\,(x_3) := \max_{u_3}\{\, c_3\, u_3 \,\} = \max_{u_{3\in\{0,1\}}}\{7.0, 7.1\} = 7$; $\mu_3^*(x_3) = 1$

La table optimale de l'étape 3 est donc :

x_3	0	1	2	3	4	5	6
$J_3\,(x_3)$	0	0	0	0	7	7	7
$\mu_3^*(x_3)$	0	0	0	0	1	1	1

Etape2

$J_2\,(x_2) = \max\limits_{u_2}\{\, 3\, u_2 + J_3\,(x_3)\} = \max\limits_{u_2}\{\, 3\, u_2 + J_3\,(x_2 - 2u_2)\}$

S.C $0 \le u_2 \le \left\lfloor \frac{x_2}{2} \right\rfloor$ *entier*

* **Pour $x_2 \in \{0, 1\}$** :

$J_2\,(x_2) = 0$; $\mu_2^*(0) = \mu_2^*(1) = 0$.

* **Pour $x_2 = 2$:**

$J_2\,(x_2) = \max_{u_{2\in\{0,1\}}}\{\, 3\, u_2 + J_3\,(x_2 - 2u_2)\} = max\{\, 3\ \times 0 + J_3\,(2), 3\ \times 1 + J_3\,(0)\}$

$= max\{0 + 0{,}3 + 0\} = 3$; $\mu_2^*(2) = 1$

* **Pour $x_2 = 3$:**

Identique au cas précédent : toujours $u_2 \in \{0{,}1\}$ et $J_3\,(3) = J_3\,(2) = 0$; $\mu_2^*(3) = 1$

* **Pour $x_2 = 4$:**

$J_2\,(x_2) = \max_{u_{2\in\{0,1,2\}}}\{\, 3\, u_2 + J_3\,(x_2 - 2u_2)\} = max\{\, 3\ \times 0 + J_3\,(4), 3\ \times 1 + J_3\,(2), 3\ \times 2 +$
$J_3\,(0)\} = \max\{0 + 7, 3 + 0, 6 + 0\} = 7$

$= max\{0 + 0{,}3 + 0\} = 3$; $\mu_2^*(4) = 0$.

Cette valeur ne change pas pour $x_2 = 5$

* **Pour $x_2 = 6$:**

$J_2\,(x_2) = \max_{u_{2\in\{0,1,2,3\}}}\{3\,u_2 + J_3\,(x_2 - 2u_2)\} = max\{3\times 0 + J_3\,(6), 3\times 1 + J_3\,(4), 3\times 2 + J_3\,(2), 3\times 3 + J_3\,(0)\} = \max\{0+7, 3+7, 6+0, 9+0\} = 10$; $\mu_2^*(6) = 1$

La table optimale de l'étape 2 est donc :

x_2	0	1	2	3	4	5	6
$J_2\,(x_2)$	0	0	3	3	7	7	10
$\mu_2^*(x_2)$	0	0	1	1	0	0	1

Etape 1

Dans cette étape l'unique problème à résoudre est :

$$J_1\,(6) = \max_{u_{1\in\{0,1,2\}}}\{5\,u_1 + J_2\,(6-3u_2)\} = max\{5\times 0 + J_2\,(6), 5\times 1 + J_2\,(3), 5\times 2 + J_2\,(0)\}$$
$$= \max\{0+10, 5+3, 10+0\} = 10$$

Il existe deux décisions optimales $\mu_1^*(6) = 0$ ou 2.

Pour chacune d'elles nous pouvons reconstruire la sélection optimale en consultant les politiques stockées dans les tables des étapes 2 et 3.

Lecture de la solution optimale :

Pour $u_1^* = 0$

L'état du système à l'étape 2, le volume encore disponible est, $x_2 = 6 - 3\mathrm{u}_1^* = 6$; la décision optimale pour cet état est $u_2^* = 1$

Le volume resté après une telle sélection est $x_3 = 6 - 2\mathrm{u}_2^* = 4$; la décision optimale pour cet état est $u_3^* = 1$.

La suite de décisions est donc {0,1,1} et correspond à un chargement formé d'un objet 2 et d'un objet 3.

La valeur de cette solution est égale à 3+7=10= $J_1\,(6)$

Pour $u_1^* = 2$: Laisser le soin de ce cas aux étudiants.

Politique obtenue {2,0,0}

3.3 Algorithme générale de programmation dynamique

Soit un problème de décisions séquentielles sur N périodes. Pour tout état initial x_1, le coût optimal $J^*(x_1)$ est égale à $J_1(x_1)$, où la fonction J_1 est donnée par la dernière étape de l'algorithme suivant :

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

Et pour k=N, …,2,1

$$J_k(x_k) \coloneqq \; \min_{u_k \in U_k(x_k)} E_{\omega_k} \left[g_k(x_k, u_k, \omega_k) + \; J_{k+1}(f_k(x_k, u_k, \omega_k))\right] \ldots..(*)$$

Si $\mu_k(x_k) \coloneqq u_k^*$ minimise la partie droite de l'équation çi-avant pour tout x_k et tout k, la politique $\pi^* = \{\mu_1^*, \mu_2^*, \dots, \mu_N^*\}$ est optimale.

Remarque

Contrairement au cas déterministe, dans un problème de décision stochastique, la politique de décision optimale ne fournit pas une suite de décisions optimales mais doit être mise en œuvre de manière séquentielle.

Exemple 2 :

Reprenons le problème de gestion de stocks

x_k: nombre d'unités disponibles au début de la période k.

u_k : le nombre d'unités commandées (reçues immédiatement) au début de la période k

ω_k : la demande aléatoire pendant la période k.

$x_{k+1} = f_k(x_k, u_k, \omega_k) = \max\{x_k + u_k - \omega_k, 0\} = [x_k + u_k - \omega_k]^+$

$x_k \in S_k = \{0,1, \dots, C\}$; C : capacité maximale du stock.

Les quantités de commande envisageable pour un niveau x_k du stock est :

$$U_k(x_k) = \{0,1, \dots, C - x_k\} \,;$$

Les coûts associés à ce système comprennent

- Un coût de réapprovisionnement : $c(u_k) = K\delta(u_k) + cu_k$

  Où :

K : coût fixe de commande ;

$$\delta(u_k) = \begin{cases} 0 & si\ u_k = 0 \\ 1 & sinon \end{cases}$$

c : coût unitaire de commande.

- Un coût de gestion :

$$r(x_k, u_k, \omega_k) = hmax\{x_k + u_k - \omega_k, 0\} - pmin\{x_k + u_k - \omega_k, 0\}$$
$$= h[x_k + u_k - \omega_k]^+ - p[x_k + u_k - \omega_k]^-$$

Où :

h : coût unitaire de stockage ;

p : coût unitaire de pénurie ;

si on considère que : $g_{N+1}(x_{N+1}) = 0\ \forall x_{N+1}$; l'espérance des coût totaux à minimiser est :

$$E\sum_{k=1}^{N}(c(u_k) + r(x_k, u_k, \omega_k))$$

<u>Question :</u> Calculer la politique optimale de réapprovisionnement pour un problème sur 3 périodes et une capacité maximale du stock C=3 , avec :

Un coût de réapprovisionnement (k=1) :

$$c(u_k) = \delta(u_k) + \frac{3}{2}u_k$$

Un coût de gestion (h=1, p= -3) :

$$r(x_k, u_k, \omega_k) = [x_k + u_k - \omega_k]^+ + 3[x_k + u_k - \omega_k]^-$$

Une loi de probabilité de la demande :

$$P[\omega_k = 0] = \frac{1}{10}, P[\omega_k = 1] = \frac{1}{5}\ P[\omega_k = 2] = \frac{3}{5} P[\omega_k = 3] = \frac{1}{10}$$

Pour k=1,2,3.

<u>Solution</u>

<u>Période4</u>

Le coût terminal étant nul : $J_4(x_4) = g_4(x_4) = 0$ pour $x_4 = 0,1,2,3$

<u>Période3</u>

$$J_3(x_3) = \min_{u_3 \in U_3(x_3)} E_{\omega_3}\left[c(u_3) + r(x_3, u_3, \omega_3) + J_4(f_3(x_3, u_3, \omega_3))\right]$$

$$= \min_{u_3 \in \{0,\ldots,3-x_3\}} \mathrm{E}_{\omega_3} \left[\delta(u_3) + \frac{3}{2} u_3 + [x_3 + u_3 - \omega_3]^+ + 3[x_3 + u_3 - \omega_3]^-\right]$$

<u>Pour $x_3 = 0$</u> , les décisions possibles sont $u_3 = 0,1,2 \; ou \; 3$

$u_3 = 0 : E[.] = \frac{1}{5} \times 3 + \frac{3}{5} \times 3 \times 2 + \frac{1}{10} \times 3 \times 3 = 5.1$ (juste ici expliquer les calculs…)

$u_3 = 1 : E[.] = 5$

$u_3 = 2 : \; E[.] = 4.7$

$u_3 = 3 : \; E[.] = 6.8$

$\mathrm{J}_3(0) = 4.7$; la quantité à commander est $\mu_3^*(0) = 2$

<u>Pour</u> $x_3 = 1$ $\;\; U_3(1) = \{0,1,2\}$

$u_3 = 0 : E[.] = 2.5 \; u_3 = 1 : E[.] = 3.2$

$u_3 = 2 : \; E[.] = 5.3$

Ainsi : $\mathrm{J}_3(1) = 2.5$ et $\mu_3^*(1) = 0$

<u>Pour</u> $x_3 = 2$ $\;\; U_3(2) = \{0,1\}$

$u_3 = 0 : E[.] = 0.7$

$u_3 = 1 : E[.] = 3.8$

Ainsi : $\mathrm{J}_3(2) = 0.7$ et $\mu_3^*(2) = 0$

<u>Pour</u> $x_3 = 3$ $\;\; U_3(3) = \{0\} \; \mu_3^*(3) = 0$ (enfin c'est la seule décision possible) et : $\mathrm{J}_3(3) = 1.3$

La table optimale de la période 3 est la suivante :

x_3	$\mu_3^*(x_3)$	$\mathrm{J}_3(x_3)$
0	2	4.7
1	0	2.5
2	0	0.7

3	0	1.3

Période2

$$J_2(x_2) = \min_{u_2 \in U_2(x_2)} E_{\omega_2}\left[c(u_2) + r(x_2, u_2, \omega_2) + J_3(f_2(x_2, u_2, \omega_2))\right]$$

$$= \min_{u_2 \in \{0,\dots,3-x_2\}} E_{\omega_2}\left[\delta(u_2) + \frac{3}{2}u_2 + [x_2 + u_2 - \omega_2]^+ + 3[x_2 + u_2 - \omega_2]^- + J_3([x_2 + u_2 - \omega_2]^+)\right]$$

Pour $x_2 = 0$

$u_2 = 0$ E[.]=9.8

$u_2 = 1$ E[.]=9.48

$u_2 = 2$ E[.]=8.56

$u_2 = 3$ E[.]=9.04

$J_2(0) = 8.56$ et la décision optimale correspondante $\mu_2^*(0) = 2$

Pour $x_2 = 1$

$u_2 = 0$ E[.]=6.98

$u_2 = 1$ E[.]=7.06

$u_2 = 2$ E[.]=7.54

$J_2(1) = 6.98$ et la décision optimale correspondante $\mu_2^*(1) = 0$

Pour $x_2 = 2$

$u_2 = 0$ E[.]=4.56

$u_2 = 1$ E[.]=6.04

$J_2(2) = 4.56$ et la décision optimale correspondante $\mu_2^*(2) = 0$

Pour $x_2 = 3$

$u_2 = 0$ E[.]=3.54

$J_2(3) = 3.54$ et la décision optimale correspondante $\mu_2^*(3) = 0$

La table optimale de la période 2 est la suivante :

x_2	$\mu_2^*(x_2)$	$J_2(x_2)$
0	2	8.56
1	0	6.98

2	0	4.56
3	0	3.54

Période1

$$J_1(x_1) = \min_{u_1 \in U_1(x_1)} E_{\omega_1}\left[c(u_1) + r(x_1, u_1, \omega_1) + J_2(f_1(x_1, u_1, \omega_1))\right]$$

$$= \min_{u_1 \in \{0,\dots,3-x_1\}} E_{\omega_1}\left[\delta(u_1) + \frac{3}{2}u_1 + [x_1 + u_1 - \omega_1]^+ + 3[x_1 + u_1 - \omega_1]^- + J_2([x_1 + u_1 - \omega_1]^+)\right]$$

La table optimale de la période 1 est la suivante :

x_1	$\mu_1^*(x_1)$	$J_1(x_1)$
0	2	12.544
1	0	10.902
2	0	8.544
3	0	7.61

3.4 Variantes et extensions

1- Les techniques de la programmation dynamique s'appliquent également aux formes de fonction coût suivantes :

- **Fonctions multiplicatives :**

L'espérance des coûts est donnée par :

$$\underset{\omega_k}{E}\left[g_{N+1}(x_{N+1}).\prod_{k=1}^{N} g_k\left(x_k, u_{k,}\omega_k\right)\right], k = 1,2, \dots, N$$

Où : $g_{N+1}(x_{N+1}) \geq 0, \forall\ x_{N+1}\ et\ g_k\left(x_k, u_{k,}\omega_k\right) \geq 0, pour\ tout\ x_k, \omega_k u_{k,}\ et\ k$

L'algorithme de la programmation dynamique est identique au cas additif, à l'exception de la formule (*) qui est remplacée par :

$$J_k(x_k) \coloneqq \min_{u_k \in U_k(x_k)} E_{\omega_k}\,[g_k(x_k, u_k, \omega_k).J_{k+1}(f_k(x_k, u_k, \omega_k))]$$

- **<u>Fonctions MINIMAX:</u>**

L'espérance des coûts est donnée par :

$$\underset{\omega_k}{\mathrm{E}}\left[\max\{g_1\left(x_k, u_{k,}\omega_k\right), g_2\left(x_k, u_{k,}\omega_k\right), \dots, g_N\left(x_N, u_{N,}\omega_N\right), g_{N+1}(x_{N+1})\}\right], k = 1,2, \dots, N$$

(*) devient: $J_k(x_k) \coloneqq \min_{u_k \in U_k(x_k)} E_{\omega_k}\,[\max\{g_k(x_k, u_k, \omega_k), J_{k+1}(f_k(x_k, u_k, \omega_k)\}]$

2. L'algorithme de la programmation dynamique est également applicable à des processus de décisions séquentielles plus généraux comportant des variables d'état et/ou de décisions multidimensionnelles (i.e vecteurs à d composantes) et/ ou continues (on procède à leurs discrétisation).

3. Il est également très facile d'inclure un facteur d'actualisation (aussi appelé facteur d'escompte) dans l'algorithme de la PD. L'objectif est alors de déterminer une politique de décision optimisant la somme (ou le produit) des coûts actualisés (escomptés).

(Voir TD pour des exemples d'application)