# La programmation dynamique et ses possibilités en calcul économique

par A. KAUFMANN (France)

*Conseiller Scientifique à la Cie des Machines BULL*

Un des instruments les plus efficaces pour rechercher l'optimum d'une fonction de valeur attachée à un phénomène économique a été présenté il y a une dizaine d'années par le mathématicien américain Richard BELLMAN; l'intérêt de cet instrument déborde d'ailleurs du domaine de l'économie et peut être estimé tout aussi important pour des recherches et évaluations en physique, en mathématiques pures et dans l'avancement des recherches technologiques.

A la base de la méthode présentée par BELLMAN se trouve un certain « principe d'optimalité » que nous préférerons d'ailleurs qualifier de « théorème d'optimalité » étant entendu qu'en mathématiques on ne doit pas utiliser des principes mais des axiomes et des théorèmes.

Ce « principe » d'optimalité est si simple qu'il parait presque trivial lorsqu'il a été bien compris, certains même le considèrent comme un truisme. Mais son importance et l'efficacité des méthodes d'optimisation séquentielle auxquelles il a donné naissance s'accentuent au fur et à mesure qu'on s'aperçoit que la vraie nature de nombreux problèmes économiques est de caractère séquentiel ; par ailleurs, certains problèmes qui n'ont pas cette nature peuvent d'une manière arbitraire être ramenés à une telle forme d'optimisation.

Historiquement, on peut faire remonter la méthode de programmation dynamique à certains travaux d'EULER, de LAGRANGE, de RAYLEIGH et de HAMILTON.

Le calcul des variations hérité de ces mathématiciens ne peut pas être dissocié de la programmation dynamique, bien que la présentation en soit toute différente.

Pour introduire pédagogiquement le principe d'optimalité, nous débuterons par un exemple et généraliserons rapidement.

Le problème du « plus court chemin », bien connu dans la théorie des graphes, est un moyen commode et surtout visuel pour faire apparaître la propriété fondamentale utilisée.

Considérons le problème suivant. Entre deux villes A et N, on se propose de construire un autoroute, qui devra passer à proximité ou à travers un certain nombre d'autres villes; il y a de nombreux points de passage possibles et un nombre assez important de tronçons peuvent être envisagés, en tenant compte de la nature du terrain et du relief. Pour chaque tronçon élémentaire, on a évalué le coût total des dépenses : coût des travaux et des ouvrages d'art, expropriations, estimation des pertes économico-sociales éventuelles des villes ou villages éloignés d'un trafic qui pouvait être profitable dans le passé.

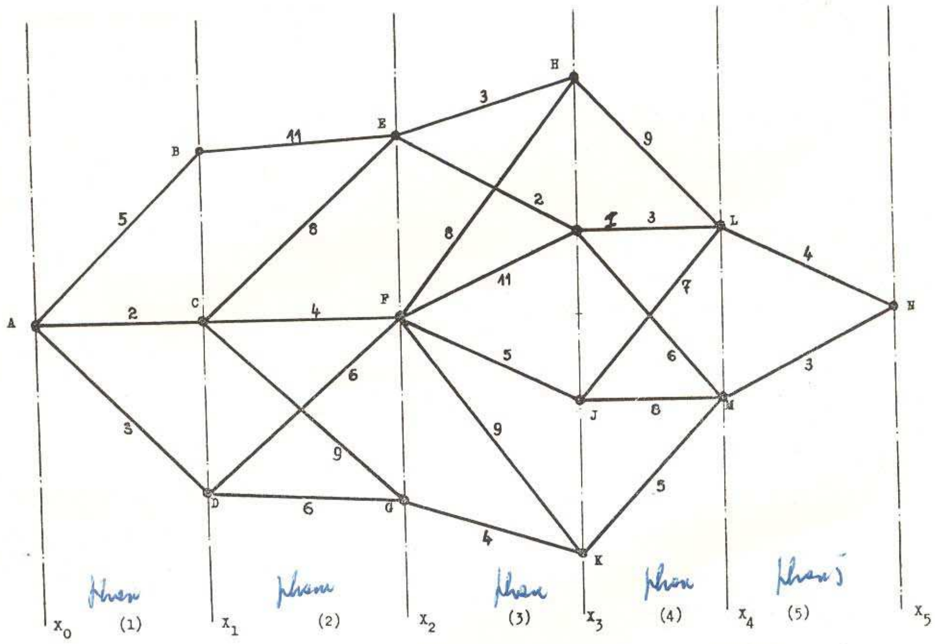


Fig. 1.

Le graphe annexé peut donner un aspect simplifié de ce genre de problème; sur ce graphe on a porté les coûts de chaque tronçon dans une unité monétaire appropriée (voir figure 1).

En phase (1) on peut choisir, à partir de A, d'aller en B, C ou D; de même, en phase (2) on peut choisir à partir de B,C ou D si cela est possible d'aller en E, F ou G; et ainsi de suite en phase (3), (4) et (5).

D.f2 Un chemin quelconque allant de A à N sera appelé une « politique ». Ainsi (ADFILN) est une « politique ». Une portion continue de chemin sera appelée une « sous-politique », ainsi (DKF), (CEIM), (HL), (CGKMN) sont des sous-politiques.

Il se trouve que (ACEILN) est la politique la moins coûteuse et donne 19; considérons toutes les sous-politiques contenues dans (ACEILN); entre les points de ce chemin, toutes les sous-politiques sont optimales, il ne peut en être autrement car il existerait un chemin meilleur de A à N, qui contiendrait une autre sous-politique meilleure, ce qui est contraire à l'hypothèse.

On est donc amené à formuler un certain truisme :

« *Une politique optimale ne peut être formée que de sous-politiques optimales* ».

BELLMAN a énoncé cette propriété autrement :

« *Une politique est optimale si, à une période donnée, quelle que soient les décisions précédentes, les décisions qui restent à prendre constituent une politique optimale en regard du résultat des décisions précédentes* ».

La définition que nous avons donnée est plus simple, mais nous devrons l'énoncer autrement lorsqu'il s'agira de processus aléatoires et nous nous rapprocherons alors de celle donnée par BELLMAN.

Appliquons au problème de l'autoroute le « principe » d'optimalité.

Soient $v_i(x_{i-1}, x_i)$ le coût d'un tronçon de la phase (i) et $F(x_0, x_1, x_2, x_3, x_4, x_5)$ le coût total de l'autoroute de A à N :

$$F(x_0, x_1, x_2, x_3, x_4, x_5) = v_1(x_0, x_1) + v_2(x_1, x_2) + v_3(x_2, x_3) + v_4(x_3\, x_4) + v_5(x_4, x_5).$$

Utilisons le principe d'optimalité, appelons $f_{1,2}(x_0, x_2)$ le coût minimal pour la portion formée par les phases (1) et (2) lorsque cette portion se termine au point $x_2$ :

$$f_{1,2}(x_0, x_2) = \text{MIN}\,[v_1(x_0, x_1) + v_2(x_1, x_2)]\,; \quad x_1 = B, C, D$$

soit :

$$\begin{aligned} f_{12}(x_0, E) &= \text{MIN}\,[5 + 11,\ 2 + 8,\ 3 + \infty] \\ &= 10 \text{ avec } x_1 = G \\ f_{12}(x_0, F) &= \text{MIN}\,[5 + \infty, 2 + 4, 3 + 6] \\ &= 6 \text{ avec } x_1 = G \\ f_{12}(x_0, G) &= \text{MIN}\,[5 + \infty, 2 + 9, 3 + 6] \\ &= 9 \text{ avec } x_1 = D \end{aligned}$$

Appelons maintenant $f_{123}(x_0, x_3)$ le coût minimal pour les tronçons 1, 2 et 3 à la fois; d'après le principe d'optimalité :

$$f_{123}(x_0, x_3) = \text{MIN}\,[f_{12}(x_0, x_2) + v_3(x_2, x_3)]\,; \quad x_2 = E, F, G$$

soit :

$$
\begin{aligned}
f_{123}\ (x_0, H) &= \text{MIN}\ [10 + 3,\ 6 + 8,\ 9 + \infty] \\
&= 13 \text{ avec } x_2 = E \\
f_{123}\ (x_0, I) &= \text{MIN}\ [10 + 2, 6 + 11, 9 + \infty] \\
&= 12 \text{ avec } x_2 = E \\
f_{123}\ (x_0, J) &= \text{MIN}\ [10 + \infty, 6 + 5, 9 + \infty] \\
&= 11 \text{ avec } x_0 = F \\
f_{123}\ (x_0, K) &= \text{MIN}\ [10 + \infty, 6 + 9, 9 + 4] \\
&= 13 \text{ avec } x_2 = G
\end{aligned}
$$

Continuons :

$$f_{1234}\ (x_0, x_4) = \text{MIN}\ [f_{123}\ (x_0, x_3) + v_4\ (x_3, x_4)]\ ;\quad x_3 = H, I, J, K$$

soit

$$
\begin{aligned}
f_{1234}\ (x_0, L) &= \text{MIN}\ [13 + 9, 12 + 3, 11 + 7, 13 + \infty] \\
&= 15 \text{ avec } x_3 = I \\
f_{1234}\ (x_0, M) &= \text{MIN}\ [13 + \infty, 12 + 6, 11 + 8, 13 + 5] \\
&= 18 \text{ avec } x_3 = I \text{ ou } x_3 = K
\end{aligned}
$$

Enfin

$$
\begin{aligned}
F^* = f_{12345}\ (x_0, x_5) &= \text{MIN}\ [f_{1234}\ (x_0, x_4) + v_5\ (x_4, x_5)]\ ;\quad x_4 = L, M \\
&= \text{MIN}\ [15 + 4, 18 + 3] \\
&= 19 \text{ avec } x_4 = L
\end{aligned}
$$

Ainsi, le chemin de coût minimal est ACEILN avec 19.

La méthode peut tout aussi bien être appliquée en commençant par la phase (5) et en remontant, ou encore en évaluant tout ensemble continu de phases et en assemblant des sous-politiques optimales ensemble.

alors Maintenant, considérons une jonction de n + 1 variables décomposable ~~de~~ la façon suivante :

$$
\begin{aligned}
F\ (x_0, x_1, x_2, x_3, \ldots\ldots, x_{n-1}, x_n) = \\
v_1\ (x_0, x_1) + v_2\ (x_1, x_2) + v_3\ (x_2, x_3) + \ldots \\
\ldots + v_n\ (x_{n-1}, x_n).
\end{aligned}
$$

Alors, la valeur optimale de F sera obtenue par récurrence, en appelant $f_{0,r}\ (x_0\ ,\ x_r)$ l'optimum pour l'ensemble des phases 0, 1, 2, ... r — 1, r consécutives, avec $F_{opt} = f_{0,n}\ (x_0\ ,\ x_n)$

$$f_{0,r}\ (x_0\ ,\ x_r) = \text{OPT}\ [f_{0,r-1}\ (x_0\ ,\ x_{r-1}) + v_r\ (x_{r-1}\ ,\ x_r)]$$

avec $x_{r-1}\ \varepsilon\ X_{r-1}\ (x_0\ ,\ x_r)$ et $f_{0,1}\ (x_0\ ,\ x_1) = v_1\ (x_0\ ,\ x_1)$.

Ces formules peuvent être aussi appliquées après adaptation des indices en remontant de n vers 0 ou par portions. Elles conviennent quelles que soient les variables $x_i$ : entières, discrètes, continues ou encore si elles représentent des vecteurs.

Bien entendu, la recherche séquentielle des optimums peut devenir très compliquée, les variables $x_i$ appartenant à des domaines qui dépendent de $x_0$ et $x_{i+1}$ et plus aussi lorsque des variables qui précèdent ou succèdent à $x_i$ telles que ($x_{i-1}$ et $x_{i+1}$), ($x_{i-2}$, $x_{i-1}$ et $x_{i+1}$, $x_{i+2}$) etc... interviennent.

On peut aussi traiter de cette manière toute fonction décomposable en phases : $F(x_0, x_1, x_2, x_3, \ldots x_{n-1}, x_n) = v_1(x_0, x_1) \times v_2(x_1, x_2) \times \ldots \times v_n(x_{n-1}, x_n)$, où le symbole $\times$ peut signifier, par exemple, une multiplication, un produit de composition, etc...

Des problèmes de stocks, d'investissements, de promotion, de localisations, de planification macro-économique peuvent être traités de cette manière, sous réserve que leur caractère soit séquentiel ou puisse y être ramené par certains artifices.

En fait, les problèmes les plus importants ne se situent pas, en général, dans un avenir certain, mais en avenir aléatoire.

Toujours à l'aide d'un graphe, donc dans des ensembles dénombrables, montrons comment le principe d'optimalité peut être étendu au cas d'un processus aléatoire, mais il sera facile de généraliser à d'autres catégories d'ensembles.

Nous allons présenter un jeu de décision et de hasard, mais la retransposition économique sera aisée.

On considère quatre cases numérotées A, B, C et D (voir figure 2).

| | |
|---|---|
| A | B |
| D | C |

Fig. 2.

Le jeu se décompose en 1 phase préalable, suivie de 3 phases. A un joueur, on donne un jeton à placer dans une case.

*Phase 0* — On indiquera au joueur dans quelle case il se trouve (condition initiale).

*Phase 1* — Etant initialement dans une certaine case, le joueur peut se déplacer selon les flèches pointillées indiquées sur la figure 3. Ainsi, étant initialement en B, il peut soit aller en A, en B ou en D.

Ce mouvement étant réalisé, le hasard intervient, une roue de loterie appropriée le renvoie dans une autre case selon les différentes lois de probabilité indiquées dans la colonne « hasard » de la phase I.

Ainsi, s'il se trouve en D, les trois événements suivants peuvent être admis : il va en A avec un gain de 3 et une probabilité de 0,5; en C avec un gain de 2 et une probabilité de 0,5; en C avec un gain de 2 et une probabilité de 0,3; en D avec un gain de 1 et une probabilité de 0,2 (voir figure 3).

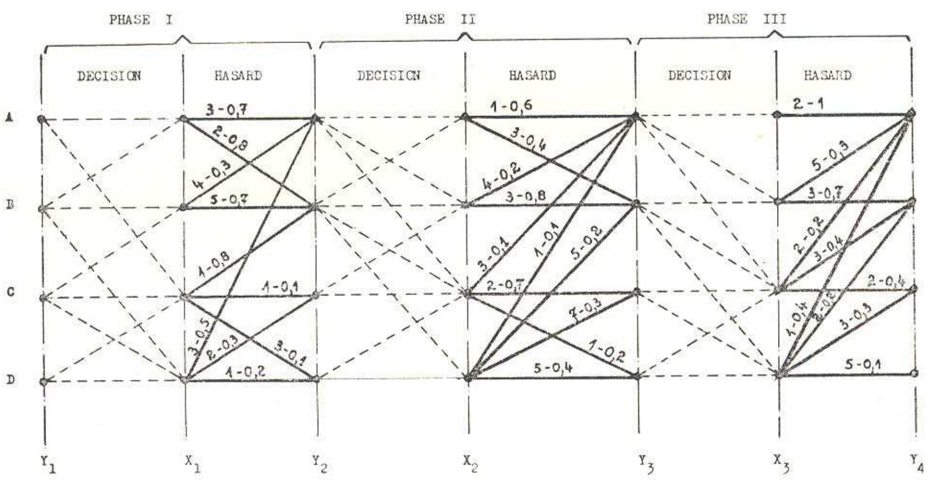


Fig. 3. — Les indications portées sur les arcs en trait pleins représentent la valeur et la probabilité de passage.

*Phases II et III* — Même processus, mais les gains et les probabilités diffèrent (le processus n'est pas stationnaire).

On répète ce jeu de hasard 100 fois; quelles sont les politiques optimales, c'est-à-dire celles qui donnent un gain maximal pour chaque situation en phase O.

En appelant $x_i$ les variables de décision et $y_j$ les variables d'état, nous opérerons comme suit en commençant par le futur, en remontant vers le passé; notons en passant qu'on ne peut faire autrement si nous admettons que le gain moyen est donné par l'espérance mathématique du gain.

*Phase III.* — Appelons $\overline{Z_3}$ $(y_3, x_3)$ l'espérance mathématique du gain quand, étant en $y_3$, on décide d'aller en $x_3$, il vient :

$\overline{Z_3}$ $(A,A) = (2)\ (1) = 2$

$\overline{Z_3}$ (A,B) = (5) (0,3) + (3) (0,7) = 3,6

$\overline{Z_3}$ (A,C) = (2) (0,2) + (3) (0,4) + (2) (0,4) = 2,4

$\overline{Z_3}$ (B,B) = (5) (0,3) + (3) (0,7) = 3,6

$\overline{Z_3}$ (B,C) = 2,4

$\overline{Z_3}$ (B,D = 2,2

$\overline{Z_3}$ (C,C) = 2,4

$\overline{Z_3}$ (C,D) = 2,2

$\overline{Z_3}$ (D,C) = 2,4

$\overline{Z_3}$ (D,D) = 2,2

En appelant $\overline{f_3}$ ($y_3$) l'espérance mathématique maximale quand on est en $y_3$, il vient :

$\overline{f_3}$ (A) = 3,6 avec $x_3$ = B

$\overline{f_3}$ (B) = 3,6 avec $x_3$ = B

$\overline{f_3}$ (C) = 2,4 avec $x_3$ = C

$\overline{f_3}$ (D) = 2,4 avec $x_3$ = C

*Phases III et II ensemble.* — Appelons $\overline{Z_2}$ ($y_2$, $x_2$) l'espérance mathématique du gain pour les phases II et III réunies quand on est en $y_2$, au début de la phase II. On appliquera le principe d'optimalité en faisant la somme du gain en phase II et de l'espérance mathématique maximale du gain en phase III pour les valeurs considérées de $y_2$, $x_2$ et $y_3$; il vient :

$\overline{Z_2}$ (A,A) = (1 + 3,6) (0,6) + (3 + 3,6) (0,4) = 5,40

$\overline{Z_2}$ (A,B) = (4 + 3,6) (0,2) + (3 + 3,6) (0,8) = 6,80

$\overline{Z_2}$ (A,C) = (3 + 3,6) (0,1) + (2 + 2,4) (0,7) + (1 + 2,4) (0,2)
= 4,42

$\overline{Z_2}$ (B,A) = 5,40

$\overline{Z_2}$ (B,B) = 6,80

$\overline{Z_2}$ (B,C) = 4,42

$\overline{Z_2}$ (B,D) = 7,96

$\overline{Z_2}$ (C,C) = 4,42

$\overline{Z_2}$ (D,C) = 4,42

En appelant $\overline{f_{3,2}}$ ($y_2$) l'espérance mathématique maximale quand on est en $y_2$, il vient :

$\overline{f}_{3,2}$ (A) = 6,80 avec $x_2$ = B

$\overline{f}_{3,2}$ (B) = 7,96 avec $x_2$ = D

$\overline{f}_{3,2}$ (C) = 6,80 avec $x_2$ = B

$\overline{f}_{3,2}$ (D) = 4,42 avec $x_2$ = C

*Phases III, II et I ensemble.* — Appelons $\overline{Z_1}$ ($y_1$, $x_1$) l'espérance mathématique du gain pour les phases III, II et I réunies quand on est en $y_1$, au début de la phase I. On appliquera encore le principe d'optimalité de la même manière :

$\overline{Z_1}$ (A,A) = (3 + 6,8) (0,7) + (2 + 7,96) (0,3) = 9,848

$\overline{Z_1}$ (A,C) = (1 + 7,96) (0,8) + (1 + 6,80) (0,1) + (3 + 4,42) (0,1)
= 8,720

$\overline{Z_1}$ (B,A) = (3 + 6,80) (0,7) + (2 + 7,96) (0,3) = 9,848

$\overline{Z_1}$ (B,B) = 12,312

$\overline{Z_1}$ (B,D) = 8,624

$\overline{Z_1}$ (C,B) = 12,312

$\overline{Z_1}$ (C,D) = 8,624

$\overline{Z_1}$ (D,C) = 8,720

$\overline{Z_1}$ (D,D) = 8,624

Finalement on aura :

$F_{MAX}$ (A) = 9,848 avec $x_1$ = A
$F_{MAX}$ (B) = 12,312 avec $x_1$ = B
$F_{MAX}$ (C) = 12,312 avec $x_1$ = B
$F_{MAX}$ (D) = 8,720 avec $x_1$ = C

Voici les politiques optimales :

si $y_1$ = A prendre $x_1$ = A

$= B$ prendre $x_1 = B$
$= C$ prendre $x_1 = B$
$= D$ prendre $x_1 = C$
si $y_2 = A$ prendre $x_2 = B$
$= B$ prendre $x_2 = D$
$= C$ prendre $x_2 = B$
$= D$ prendre $x_2 = C$
si $y_3 = A$ prendre $x_3 = B$
$= B$ prendre $x_3 = B$
$= C$ prendre $x_3 = C$
$= D$ prendre $x_3 = C$

Il serait commode de convertir le jeu que nous venons d'étudier en problème de stocks ou d'entretien. En fait, dans le cas aléatoire, les problèmes pratiques se présentent sous la forme de phases : décision - hasard (D, H) ou hasard - décision (H, D), ou encore des phases plus complexes, par exemple (D, D, H, D, H).

Dans le cas aléatoire, le principe d'optimalité sous la forme où nous l'avons énoncé et qui différait de celle de BELLMAN, doit être modifié ici et s'énoncer :

« *Une sous-politique optimale de la phase N (la dernière) à la phase (N - n) ne peut être formée que par une sous-politique optimale de N à (N — n + 1)* »

Indiquons que tout processus aléatoire séquentiel (D,H), (H,D) ou plus compliqué, dont les variables d'état et de décision d'une phase ne dépendent que de celles de la phase précédente peut être ramené à une chaîne de MARKOV multiple avec revenu (selon la dénomination de HOWARD).

A titre indicatif, donnons la description d'une chaîne de MARKOV multiple avec revenu dans le cas stationnaire et montrons comment s'applique le principe d'optimalité.

Considérons un processus aléatoire de décision discret tel que, à tout changement d'état $E_i$ vers $E_j$, $i, j = 1, 2, \ldots, M$ on peut faire correspondre une probabilité $p_{ij}^{(r)}$. Pour chaque état $E_i$, on considère un ensemble de $m_i$ vecteurs stochastiques $[p_i^{(r)}] = [p_{i1}^{(r)}, p_{i2}^{(r)}, \ldots, p_{iM}^{(r)}]$ ; le choix d'un tel vecteur d'indice r parmi les $m_i$ vecteurs est libre et constitue une décision. A toute probabilité $p_{ij}^{(r)}$ on associe une valeur ou revenu $R_{ij}^{(r)}$ qui est un nombre réel; on forme ainsi des vecteurs de revenu $[R_i^{(r)}] = [R_{i1}^{(r)}, R_{i2}^{(r)}, \ldots, R_{iM}^{(r)}]$, qui seront associés aux vecteurs stochastiques correspondants.

On appelle « politique » une séquence de décisions sur un ensemble de périodes (phases) jointives. Considérons N transitions successives effectuées aux dates 0,1,2 ..., N-1; on demande quelle est la ou les politiques optimales, c'est-à-dire qui rendent maximale (ou tout aussi bien minimale dans un autre cas) la somme des valeurs des N transitions.

Appelons $q_i^{(r)}$ l'espérance mathématique de la valeur d'une transition quand on est dans l'état $E_i$ à une date quelconque n et que l'on a choisi le vecteur stochastique $[p_i^{(r)}]$. On a :

$$q_i^{(r)} = \sum_{j=1}^{M} p_{ij}^{(r)} R_{ij}^{(r)}, \quad i = 1, 2, \ldots, M.$$

Soit $v_i$ (N — n, N) l'espérance mathématique de la valeur totale sur n transitions depuis la date N — n jusqu'à la date N lorsque à la date N — n le système est dans l'état $E_i$ ; alors, en vertu du principe d'optimalité :

$$v_i (N - n, N) = \underset{r=1,2,\ldots n}{\mathrm{MAX}} \left[q_i^{(r)} + \sum_{j=1}^{M} p_{ij}^{(r)} v_j (N - n + 1, N)\right]$$
$$i = 1, 2, \ldots, M.$$

Pour le mathématicien, d'une part, et pour l'économiste ou analyste, d'autre part, de bien belles questions se posent à propos de ces chaînes de MARKOV multiples dans le cas où elles sont stationnaires : convergence (ergodisme), décomposition de l'ensemble des états en classes d'équivalence, de connexité, approximation dans l'espace des politiques.

Des procédés nouveaux ont été récemment introduits pour approcher rapidement la ou les politiques optimales en utilisant des calculateurs électroniques.

Ces derniers s'avèrent indispensables, pratiquement dès qu'on s'intéresse à des problèmes réels d'aspect économique; c'est le cas, par exemple, des problèmes d'entretien et remplacement, de nombreux problèmes de production et stocks, de démographie économique, etc...

Quand on s'intéresse à la programmation dynamique, on ne peut oublier certains prolongements qui s'imposent. En tout premier lieu, les problèmes à anticipation, à apprentissage ou à adaptation. Il s'agit, en fait, de serrer la réalité de plus près, en acceptant de reconnaître combien il est arbitraire de se donner a priori un certain horizon économique, au lieu d'étudier la façon dont l'information issue du passé se disperse au fur et à mesure qu'on anticipe plus loin, mais cette étude n'est pas commode dans l'état actuel de nos connaissances.

Les récents travaux de BELLMAN et de son équipe, en ce qui concerne les systèmes à apprentissage (ou adaptation) sur les contrôles séquentiels peuvent être avec bonheur étendus à des processus économiques dans l'entreprise ou dans des structures plus larges.

Il est très excitant de reconnaître combien sont analogues les préoccupations des ingénieurs en fusées cosmiques et celles des économistes d'entreprise; le déploiement se fait par étage, phase par phase, avec correction ou adaptation.

La programmation dynamique est la méthode de calcul de l'ingénieur en fusées, elle doit former une partie importante dans un proche futur, de la liste des instruments permettant la préparation de la décision et son adaptation aux perturbations de l'environnement.