G. TH. GUILBAUD

Programmes dynamiques et programmes linéaires Note sur un modèle de Richard Bellman

Cahiers du Bureau universitaire de recherche opérationnelle. Série Recherche, tome 2 (1957), p. 37-41

<http://www.numdam.org/item?id=BURO_1957__2__37_0>

PROGRAMMES DYNAMIQUES ET PROGRAMMES LINÉAIRES NOTE SUR UN MODÈLE DE RICHARD BELLMAN

par

G. Th. GUILBAUD

On dispose d'une quantité <u>donnée</u> (stock ou capital initial), soit : a_0. On peut l'utiliser de plusieurs façons, mais les rendements sont différents. Mais on peut aussi diviser la quantité a en plusieurs parts, chacune étant affectée à l'une des utilisations possibles.

Raisonnons sur le cas le plus simple où il n'existe que deux modes d'emploi. On effectuera d'abord un partage :

$$a_0 = a'_0 + a''_0$$

(Les quantités a'_0 et a''_0 sont toutes deux non négatives).

Le rendement global est supposé somme des rendements partiels, soit :

$$r_0 = f'(a'_0) + f''(a''_0)$$

les deux fonctions f' et f'' qui définissent les rendements des deux modes d'emploi sont données.

Après un tel emploi le capital redevient disponible, mais avec une perte (ou usure) ; le capital initial étant :

$$a_0 = a'_0 + a''_0$$

on posera que le capital final sera

$$a_1 = g'(a'_0) + g''(a''_0)$$

les deux fonctions g' et g'' caractérisant l'usure. Par exemple on pourra prendre deux fonctions linéaires telles que :

$$g(a) = g \cdot a$$

dans laquelle le coefficient g est inférieur à l'unité (la différence 1-g étant le taux d'usure). Mais l'hypothèse de proportionnalité n'est pas indispensable, et dans ce qui va suivre, nous conserverons la formulation générale, avec des fonctions g quelconques.

A côté de diverses interprétations économiques, on peut songer à une illustration "stratégique" : le stock initial a_o désignant les forces disponibles, qu'il s'agit de répartir entre plusieurs théâtres d'opérations. Les rendements $f'(a'), f''(a'')$ pourront désigner, par exemple, les pertes ennemies.

II

Ainsi donc, partant de a_o, on obtient, pour un partage donné un rendement

$$r_o = f'(a'_o) + f''(a''_o)$$

et un nouveau capital

$$a_1 = g'(a'_o) + g''(a''_o)$$

On va pouvoir recommencer l'opération, c'est-à-dire effectuer un nouveau partage

$$a_1 = a'_1 + a''_1$$

qui à son tour donnera

$$r_1 = f'(a'_1) + f''(a''_1)$$

$$a_2 = g'(a'_1) + g''(a''_1)$$

et ainsi de suite.

On se trouve donc en présence d'une séquence de décisions de partage et l'on se propose de déterminer la meilleure. On pourrait, par exemple, fixer le nombre d'opérations successives et définir l'optimum par le maximum de la somme de tous les rendements :

$$r_o + r_1 + r_2 + \dots + r_n$$

On pourrait aussi introduire dans le critère, une "valeur" du résidu final a_{n+1}, on examinera plus loin ce perfectionnement du modèle, ainsi que quelques autres.

III

Plaçons-nous au début de la dernière étape. Soit x le stock restant (inconnu en ce moment de notre raisonnement). On peut calculer le rendement de la dernière décision, caractérisée par le partage :

$$x = (x-y) + (y)$$

Le rendement est égal à :

$$f'(x-y) + f''(y)$$

Et ce qu'on peut faire de mieux, c'est choisir de telle façon que cette somme soit maximum. Posons

$$E_1(x) = \underset{(y)}{\mathrm{Max}}\left[f'(x-y) + f''(y)\right]$$

étant entendu que :

$$0 \leqslant y \leqslant x$$

On a donc obtenu une règle (partielle) d'action pour la dernière étape.

Prenons maintenant l'avant-dernière : si le stock restant est u et si l'on fait le partage :

$$u = (u-v) + (v)$$

le rendement sera

$$f'(u-v) + f''(v)$$

et le résidu

$$x = g'(u-v) + g''(v)$$

Si l'on applique la règle partielle déjà trouvée, le rendement des deux dernières étapes sera :

$$f'(u-v) + f''(v) + E_1\left[g'(u-v) + g''(v)\right]$$

On cherchera à le rendre maximum, $\underline{u}$ étant donné, et v choisi entre les limites :

$$0 \leqslant v \leqslant u$$

On posera :

$$E_2(u) = \mathrm{Max}\left[f'(u-v) + f''(v) + E_1\left[g'(u-v) + g''(v)\right]\right]$$

On va continuer ainsi par récurrence, en posant :

$$E_{h+1}(z) = \underset{(0 \leqslant t \leqslant z)}{\mathrm{Max}}\left[f'(z-t) + f''(t) + E_h\left[g'(z-t) + g''(t)\right]\right]$$

ce qui donne la valeur d'un stock z lorsqu'il reste encore (h+1) étapes, et lorsqu'on décide d'effectuer de la meilleure façon possible tous les choix ultérieurs.

On notera la similitude de cette méthode avec la définition générale de l'espérance mathématique (Règle des Partis de Pascal et Théorème de Zermelo pour les jeux). Cf. Leçons sur les éléments principaux de la théorie mathématique des Jeux, leçon II, pages 1-7, dans "Stratégie et décisions économiques", Paris, C.N.R.S., 1954.

IV

Traitons complètement un cas simple, celui dans lequel les fonctions f et g sont linéaires. Il n'est évidemment pas nécessaire de se limiter à deux façons seulement d'employer le stock existant. On posera donc :

$$f^{(i)}(x_i) = f_i x_i \qquad g^{(i)}(x_i) = g_i x_i$$

et on doit résoudre l'équation fonctionnelle de récurrence :

$$E_{h+1}(x) = \mathrm{Max}\left[\Sigma f_i x_i + E_h(\Sigma g_i x_i)\right]$$

$$\Sigma x_i = x \quad , \quad x_i \geqslant 0$$

à partir de :

$$E_1(x) = \mathrm{Max}(\Sigma f_i x_i) \quad \text{pour} \quad \Sigma x_i = x \quad \text{et} \quad x_i \geqslant 0$$

il est clair que :

$$E_1(x) = e_1 x$$

avec :

$$e_1 = \operatorname*{Max}_{(i)} (f_i)$$

On peut montrer, par récurrence, que l'on a d'une façon analogue :

$$E_h(x) = e_h x$$

en effet, si cela est vrai pour h, on a

$$\Sigma f_i x_i + E_h(\Sigma g_i x_i) = \Sigma (f_i + e_h g_i) x_i$$

d'où :

$$E_{h+1} = \max_{(i)} \Sigma (f_i + e_h g_i) x_i \qquad \text{donc :} \quad E_{h+1} = e_{h+1} x$$

avec

$$e_{h+1} = \operatorname*{Max}_{(i)} (f_i + e_h g_i)$$

Pour déterminer la solution il suffira de tracer des droites :

$$y = f_i + g_i x$$

et d'en prendre l'enveloppe supérieure, que nous désignerons par :

$$y = L(x) = \operatorname*{Max}_{(i)} (f_i + x g_i)$$

Cette fonction, qui est représentée par une ligne polygonale, permet de calculer la suite des coefficients e_h, par la loi :

$$e_o = o \quad , \qquad e_{h+1} = L(e_h)$$

Comme tous les coefficients g_i sont inférieurs à l'unité, la pente de $L(x)$ est toujours inférieure à l'unité et par conséquent la suite des e_h a une limite lorsque h augmente indéfiniment.

Il en résulte, dans ce cas, que lorsque la suite des opérateurs est très longue, le résultat est à peu près indépendant de sa longueur. Ce résultat à très long terme est donné par la solution de :

$$x = L(x)$$

c'est-à-dire qu'on calculera toutes les solutions de :

$$x = f_i + g_i x$$

soit :

$$x = f_i : (1 - g_i)$$

et qu'on choisira la plus grande :

$$e = \max_{(i)} \left[f_i : (1 - g_i) \right] = f_p : (1 - g_p)$$

Alors :

1°) Le rendement maximum qu'on puisse tirer d'un stock initial x, à condition d'avoir le temps, sera égal à e.x.

2°) La politique la meilleure consiste à concentrer l'emploi de tout le stock disponible sur l'emploi numéroté (p).

V

Il peut être intéressant de comparer la méthode précédente aux méthodes classiques utilisées pour résoudre les problèmes de programmation linéaire.

Désignons par x le stock disponible à l'étape (t) ; on fait le partage :

$$x_t = x_t^1 + x_t^2 + \dots + x_t^m = \Sigma\ x_t^i$$

d'où le nouveau stock :

$$x_{t-1} = \Sigma\ g^i\ x_t^i$$

et le nouveau partage :

$$\Sigma\ g^i\ x_t^i = \Sigma\ x_{t+1}^i \qquad (1)$$

Pour le départ, on a :

$$x_0^i = a \text{ (donné)} \qquad (2)$$

Le rendement global est :

$$R = \Sigma\ f^i\ x_0^i + \Sigma\ f^i\ x_1^i + \dots + \Sigma^i f_n^i\ x_n^i \qquad (3)$$

On doit donc chercher le maximum de (3) en choisissant les x_t^i liés par les relations (1) et (2) et assujettis de plus à la condition : $x_t^i \geqslant 0$.

Utilisons la dualité. Prenons comme multiplicateurs de (1) des nombres e_{t+1},et e_0 pour (2). On doit avoir

$$e_t \geqslant f^i + g^i \cdot e_{t+1} \qquad (t = 0, 1, \dots, n-1)$$

et chercher le minimum de : $e_0 a$

c'est-à-dire de e_0.

On retrouve donc exactement la procédure précédente.