INSTITUT DE STATISTIQUE
DE L'UNIVERSITÉ DE PARIS

CAHIERS DU BUREAU UNIVERSITAIRE DE RECHERCHE OPÉRATIONNELLE

Cahier n° 2

Marcel-Paul SCHUTZENBERGER

SUR UNE GÉNÉRALISATION DE L'INÉGALITÉ MINIMAX

Marc BARBUT

MÉTHODES RECURRENTES DANS LES PROBLÊMES
DE RENOUVELLEMENT DE STOCK

Hachiro AKAMA

UN ASPECT DE LA PROGRAMMATION DYNAMIQUE

G. Th. GUILBAUD

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

1957

PARIS
INSTITUT HENRI POINCARÉ

SUR UNE GÉNÉRALISATION DE L'INÉGALITÉ MINIMAX

par

Marcel-Paul SCHÜTZENBERGER

INTRODUCTION

Soit donné un ensemble : $A = (a_j^i)$ $(i = 1, 2, \ldots, n ; j = 1, 2, \ldots, m)$ de $(n.m)$ valeurs numériques .

L'inégalité dite "minimax" :

$$\operatorname*{Min}_i \operatorname*{Max}_j a_j^i \geqslant \operatorname*{Max}_j \operatorname*{Min}_i a_j^i$$

joue un rôle fondamental dans la théorie des jeux et la recherche opérationnelle. Or cette inégalité ne fait appel, pour sa démonstration, qu'aux propriétés les plus élémentaires des opérations associées à la relation $(\leqslant)$ entre grandeurs. Il peut donc présenter un certain intérêt de savoir dans quelle mesure elle se généralise quand on substitue aux opérations "Min" et "Max" des opérations consistant à prendre non plus seulement le premier (min) et le dernier (max) d'un ensemble d'éléments rangés dans l'ordre des grandeurs croissantes mais, plus généralement, le "p-ième". Ce que nous noterons par $\overset{(p)}{M}$. On a : Min = $\overset{(1)}{M}$; Max = $\overset{(N)}{M}$ (si N est le nombre des éléments de l'ensemble).

Comme les a_j^i n'interviennent que par leur grandeur relative il est sans importance de les remplacer par les n.m valeurs $1, 2, \ldots, n.m$ et les raisonnements seront sans doute plus faciles à suivre si nous examinons d'abord l'exemple numérique ci-dessous.

EXEMPLE

$$\text{Soit } A = \begin{vmatrix} 20 & 14 & 1 & 4 & 17 \\ 6 & 10 & 19 & 18 & 16 \\ 15 & 3 & 7 & 13 & 5 \\ 2 & 8 & 9 & 11 & 12 \end{vmatrix}$$

A partir de A construisons les deux tableaux B = Col. A et C = Lig. A obtenues en réarrangeant par ordre de grandeurs les éléments de chaque colonne (ligne) :

$$
B = \begin{vmatrix} 2 & 3 & 1 & 4 & 5 \\ 6 & 8 & 7 & 11 & 12 \\ 15 & 10 & 9 & 13 & 16 \\ 20 & 14 & 19 & 18 & 17 \end{vmatrix} \qquad C = \begin{vmatrix} 1 & 4 & 14 & 17 & 20 \\ 6 & 10 & 16 & 18 & 19 \\ 3 & 5 & 7 & 13 & 15 \\ 2 & 8 & 9 & 11 & 12 \end{vmatrix}
$$

Réarrangeons maintenant B par ligne et C par colonne. Nous obtenons

$$
D = \text{Lig. } B = \text{Lig. Col. } A = \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 11 & 12 \\ 9 & 10 & 13 & 15 & 16 \\ 14 & 17 & 18 & 19 & 20 \end{vmatrix}
$$

$$
E = \text{Col. } C = \text{Col. Lig. } A = \begin{vmatrix} 1 & 4 & 7 & 11 & 12 \\ 2 & 5 & 9 & 13 & 15 \\ 3 & 8 & 14 & 17 & 19 \\ 6 & 10 & 16 & 18 & 20 \end{vmatrix}
$$

Désignant par des lettres minuscules les éléments des tableaux correspondants il résulte de la construction même et selon des notations évidentes :

(1) $\quad b_k^i = \overset{k}{M}\; a_j^i \; ; \; c_j^h = \overset{h}{M}\; a_j^i \; ; \; d_k^h = \overset{h}{M}_i\, b_k^i = \overset{h}{M}_i\, \overset{k}{M}_j\, a_j^i \; ; \; e_k^h = \overset{k}{M}_j\; c_j^h = \overset{k}{M}_j\, \overset{h}{M}_i\, a_j^i$

Il apparaît sur l'exemple choisi un phénomène remarquable : les éléments de D situés dans l'angle supérieur droit sont systématiquement plus petits que les éléments correspondants de E et l'inverse est vrai pour l'angle inférieur gauche.

En particulier on retrouve bien :

$$\underset{j}{\mathrm{Max}}\ \underset{i}{\mathrm{Min}}\ a_j^i \leqslant \underset{i}{\mathrm{Min}}\ \underset{j}{\mathrm{Max}}\ a_j^i$$

(soit ici $5 < 12$)

Le lecteur pourra d'ailleurs vérifier en prenant au hasard des valeurs a_j^i que ces particularités n'ont rien d'exceptionnel et le problème se pose donc de savoir quelles sont les conditions nécessaires et suffisantes pour que

(2) $$\underset{(i)}{\overset{(h)}{M}}\ \underset{(j)}{\overset{(k)}{M}}\ a_j^i \leqslant \underset{(j)}{\overset{(k)}{M}}\ \underset{(i)}{\overset{(h)}{M}}\ a_j^i$$

CALCUL DE $\overset{h}{M}_i$

Il est bien connu que, N quantités numériques u^i étant données, la fonction latticielle symétrique $\overset{(h)}{M}_i u_i$ est obtenue aussi bien par l'une que par l'autre des deux formules suivantes :

(3) $\overset{(\alpha)}{M}_i u^i$ = "Somme" de tous les "produits" $(N-\alpha+1)$ à $(N-\alpha+1)$ des u^i

(3') $\overset{\alpha}{M}_i u^i$ = "Produit" de toutes les "sommes" α à α des u^i.

Les mots "Somme" et "Produits" devant être entendus dans leur sens latticiel, c'est-à-dire que :

- "Somme" des éléments d'un ensemble X = les plus grands éléments de X,
- "Produit" des éléments de X = le plus petit élément de X.

Par exemple : Si $[x \leqslant y \leqslant z \leqslant t] = X$ on a :

$\overset{1}{M}X = xyzt = x = \text{Min } X =$ Produit de x, y, z et t

$\overset{2}{M}X = xyz + xyt + xzt + yzt = x + x + x + y = y$

ou $= (x+y)(x+z)(x+t)(y+z)(y+t)(z+t) = yztyyz = y$

$\overset{3}{M}X = xy + xz + xt + yz + yt + zt = x + x + xy + y + z = z$

ou $(x+y+z)(x+y+t)(x+z+t)(y+z+t) = zttt = z$

$\overset{4}{M}X = x + y + z + t = t = \text{Max } X =$ "Somme" de x, y, z et t

On sait que les opérations "Somme" et "Produit" sont à la fois associatives, commutatives, absorbantes ($x(x+y) = x + xy = x$) et distributives ($x(x+z) = xy + xz$; $(x+y)(x+z) = x+yz$) et on observera que la définition (3) ou (3') donne un sens à $\overset{\alpha}{M}$ même quand plusieurs des quantités considérées sont égales entre elles. Par exemple, si $x = y$: $\overset{1}{M}X = \overset{2}{M}X$ = la valeur commune de x et de y.

CALCUL DES $\overset{\alpha}{M}_i \overset{\beta}{M}_j$

D'après (3) on a pour un β fixé :

$d^\alpha_\beta = \overset{\alpha}{M}_i \overset{\beta}{M}_j a^i_j$ = "Somme" des "Produits" $(n-\alpha+1)$ à $(n-\alpha+1)$ des b^i_β

et encore d'après (3) pour un i fixé :

b^i_β = "Somme" des "Produits" $(n-\beta+1)$ à $(n-\beta+1)$ des a^i_j

Donc, en vertu de la distributivité des deux opérations l'une sur l'autre et compte tenu de leur associativité :

d^{α}_{β} = "Somme" de tous les "Produits" d_k de tous les éléments appartenant à un ensemble D_k obtenu lui-même en choisissant un ensemble A_k de $n-\alpha+1$ colonnes de A et dans chacune de celles-ci $m-\beta+1$ éléments a^i_j.

k parcourt un ensemble K comprenant $\left[\begin{smallmatrix} n \\ n-\alpha+1 \end{smallmatrix}\right] \left[\begin{smallmatrix} m \\ m-\beta+1 \end{smallmatrix}\right]$ indices distincts et :

$$(4) \qquad d^{\alpha}_{\beta} = \sum_{k \in K} d_k \qquad \text{où} \qquad d_k = \prod_{a^i_j \in D_K} a^i_j$$

$$(4') \qquad e^{\alpha}_{\beta} = \prod_{\ell \in L} e_\ell \qquad \text{où} \qquad e_\ell = \sum_{a^i_j \in E_\ell} a^i_j$$

E_ℓ étant obtenu en choisissant α éléments a^i_j dans chacune des β lignes d'un ensemble B_ℓ ($\ell \in L$).

Pour la suite, nous aurons avantage à introduire encore les notations suivantes relatives à un $k \in K$ et un $\ell \in L$ fixés :

$A_1 = a^i_j$	tel que:	$i \in A_k$	$j \in B_\ell$
$A_2 = a^i_j$	"	$i \in A_k$	$j \notin B_\ell$
$A_3 = a^i_j$	"	$i \notin A_k$	$j \in B_\ell$

Donc : $D_k \subset A_1 \cup A_2$; $E_\ell \subset A_1 \cup A_3$

CONDITIONS DE VALIDITÉ IDENTIQUE DE L'INÉGALITÉ (2)

Il est bien connu que quelque soient x, y et z on a identiquement : $xy \leqslant x \leqslant x+z$ les opérations ayant évidemment leur sens latticiel. En particulier, quelque soient les u^i[3] implique $\overset{\alpha}{M}_i u^i \leqslant M_i u^i$ et, d'autre part, si $u^i \leqslant v^i$ pour tout i, alors $\overset{\alpha}{M}_i u^i \leqslant M_i v^i$.

Donc si $\alpha \leqslant \alpha'$ et $\beta \leqslant \beta'$:

$$\overset{\alpha}{M}_i \overset{\beta}{M}_j a^i_j \leqslant \overset{\alpha'}{M}_i \overset{\beta'}{M}_j a^i_j$$

ce qui s'interprète immédiatement sur les tableaux D et E. Revenons à l'inégalité (2). Celle-ci s'écrit :

$$d^{\alpha}_{\beta} = \sum d_k \leqslant e^{\alpha}_{\beta} = \prod e_\ell$$

et est donc satisfaite sauf si $e_{i_1} \leqslant d_{k_1}$ pour au moins un couple d'indices $k_1 \in K$ et $l_1 \in L$. Mais si les ensembles D_k et E_i possédaient un élément $a^{i_1}_{j_1}$

commun on aurait $d_k \leqslant a_{j_1}^{i_1}\, e_1$ et, par conséquent, (2) ne peut se trouver en défaut que si les deux conditions suivantes sont satisfaites :

(5) Il existe au moins un couple (D_{k_1}, E_{l_1}) tel que $D_{k_1} \cap E_{l_1} = 0$

(6) $d_{k_1} = \text{Min } D_{k_1}, e_{l_1} = \text{Max}_{l_1}$

En particulier, puisque D_k contient $(n-\alpha+1)(m-\beta+1)$ éléments et $E, \alpha\beta$, il sera possible de satisfaire à (5) si :

(7) $(n-\alpha+1)(m-\beta+1) + \alpha\beta > n + (n-\alpha+1)\, m - \beta(n-\alpha+1)$ c'est-à-dire si :

(7') $(\beta-1)(n-\alpha) > 1$

puisque le membre de droite de (7) est le nombre des cases de l'ensemble $A_1 \cup A_2 \cup A_3$. On trouve ainsi une généralisation de l'inégalité Minimax :

(2) $\text{Max}_i \overset{\beta}{M}_j\, a_j^i \leqslant \overset{\beta}{M}_j\, \text{Max}_i \leqslant a_j^i$ pour tout β $(1 \leqslant \beta \leqslant m)$

$\overset{\alpha}{M}_i\, \text{Min}_j\, a_j^i \leqslant \text{Min}_j\, \overset{\alpha}{M}_i\, a_j^i$ pour tout α $(1 \leqslant \alpha \leqslant n)$.

Supposons maintenant que $\beta \neq 1$ et $\alpha \neq n$ et montrons que l'on peut satisfaire à (5), A_k et B_ℓ étant fixés.

En effet, dans chaque colonne de A_k, $m-\beta+1$ éléments a_j^i seulement figurent dans D_k. On peut donc choisir cet ensemble de telle façon que $m-\beta+1-(m-\beta)$ = un seul a_j^i par colonne se trouve dans A_1 et de même pour les lignes. D'où le résultat puisque A_1 étant un rectangle qui comporte au moins deux lignes et deux colonnes il est possible de faire en sorte que ces deux derniers ensembles de a_j^i soient disjoints.

Comme (6) peut toujours être satisfaite, il en résulte que (2) est le seul cas où (2) soit identiquement vérifié. <u>Exemple</u> : $n = m = 4$. Pour le tableau suivant on a :

$$8 = \overset{3}{M}_i\, \overset{2}{M} a_j^i > \overset{2}{M}_j\, \overset{3}{M}_i\, a_j^i = 7$$

(A_k : les colonnes 2 et 4 ; B_{l_1} les lignes 2 et 3)

$$A = \begin{vmatrix} 11 & 8 & 16 & 13 \\ 3 & 14 & 1 & 7 \\ 6 & 2 & 5 & 9 \\ 4 & 12 & 10 & 15 \end{vmatrix} \quad B = \begin{vmatrix} 3 & 2 & 1 & 7 \\ 4 & 8 & 5 & 9 \\ 6 & 12 & 10 & 13 \\ 11 & 14 & 16 & 15 \end{vmatrix} \quad C = \begin{vmatrix} 8 & 11 & 13 & 16 \\ 1 & 3 & 7 & 14 \\ 2 & 5 & 8 & 9 \\ 14 & 10 & 12 & 15 \end{vmatrix}$$

$$D = \begin{vmatrix} 1 & 2 & 3 & 7 \\ 4 & 5 & 0 & 9 \\ 6 & 10 & 12 & 13 \\ 11 & 14 & 15 & 16 \end{vmatrix} \quad E = \begin{vmatrix} 1 & 3 & 6 & 9 \\ 2 & 5 & 0 & 14 \\ 4 & 10 & 12 & 15 \\ 8 & 11 & 13 & 16 \end{vmatrix}$$

Validité limite de (2)

Il apparaît cependant vraisemblable pour des raisons de symétrie que si β/m est plus petit que α/n (2) doit être "généralement vraie" dans un sens qui reste à préciser. Nous le ferons ici très simplement sous les hypothèses suivantes :

1°- $n = m$ tend vers l'infini

2°- $\beta = n-\alpha+1$ est l'entier le plus voisin de λn ou λ est une constante finie $\ll 1/2$

3°) Les $(n^2)!$ permutations possibles des a_j^i sont équiprobables.

$\overset{\beta}{M}$ est donc le "λ-ième quantile inférieur" $\overset{\lambda}{M}$ et nous chercherons à prouver que si λ est assez petit, zéro est la limite quand $n \longrightarrow \infty$ de la probabilité P que :

$$(2°) \quad \overset{1-\lambda}{M_i}\ \overset{\lambda}{M_j}\ a_j^i > M_j\ \overset{1-\lambda}{M_i}\ a_j^i$$

quand les a_j^i sont tirés au sort indépendamment et selon une seule et même loi de probabilité.

Appelons $E_{k,\ell}$ l'événement consistant en ce que (6) soit vérifié pour un couple (D_k, E_ℓ). On a :

(8) $P < \sum Pr(E_{k,\ell})$ où la sommation est étendue à tous les couples (D_k, E_ℓ) tels que $D_k \cap E_\ell = \emptyset$. En raison de la symétrie (8) peut s'écrire :

$$(8') \quad P \left[{n \atop \lambda n}\right]^2 \times Q \times P_0 \qquad \text{où :}$$

$P_0 = Pr(E_{k,\ell})$ est indépendant des indices k et l.

Q est le nombre des façons de choisir D_k et E_ℓ disjoints quand A_k et B_ℓ sont fixés.

$\left[{n \atop n\lambda}\right]^2$ = est le nombre des façons de choisir A_k et B_ℓ.

La seule difficulté réside dans le calcul de Q. En effet, D_k et E_ℓ étant fixés et contenant chacun $n' = \lambda n\,(n-\lambda+1)$ éléments la probabilité pour que (6) soit vraie est :

$$(9) \quad P_0 = \left[{2n' \atop n}\right]^{-1} = \exp\left(-n^2(\lambda-\lambda^2)\, 2 \operatorname{Log} 2 + O(n^2)\right).$$

Au lieu de calculer Q nous calculerons Q' qui en est une estimation grossière pour λ voisin de $1/2$ mais suffisante pour λ petit.

Q' = Nombre des façons de prendre deux ensembles D' et E' de n' éléments chacun de telle sorte que $D' \subset A_1 \cup A_2$; $E' \subset A_1 \cup A_3$ et $D' \cap E' = \emptyset$.

Supposons qu'en outre il soit fixé que X éléments de D' et Y éléments de E′ sont dans A_1 qui contient $\lambda^2 n^2$ cases, on a :

$$Q'ky = \frac{(\lambda^2 n^2)!}{X!\,Y!\,(\lambda^2 n^2 - X - Y)!} \begin{bmatrix} n^2 (\lambda - \lambda^2) \\ n' - X \end{bmatrix} \begin{bmatrix} n^2 (\lambda - \lambda^2) \\ n' - Y \end{bmatrix} \quad \text{Donc :}$$

$Q' = \sum Q'_{x,y}$ où la sommation est étendue à toutes les valeurs de X et de Y telles que $X+Y \leqslant \lambda^2 n^2$. Posons

$$X = x\,\lambda^2 n^2 \quad \text{et} \quad Y = y\,\lambda^2 n^2$$

$$H(u) = u \log 1/u + (1-u) \operatorname{Log} 1/(1-u) ;$$

$$H(x,y) = x \operatorname{Log} 1/x + y \operatorname{Log} 1/y + (1-x-y) \operatorname{Log} 1/(1-x-y)$$

il vient

$$\operatorname{Log} Q'_{xy} = n^2 \lambda^2 H(x,y) + n^2(\lambda-\lambda^2)(H(\tfrac{x\lambda}{1-\lambda}) + H(\tfrac{y\lambda}{1-\lambda})) + o(n^2)$$

$$= n^2(\lambda-\lambda^2)\, F_\lambda(x,y) + o(n^2)$$

Q'_{xy} étant exponentiellement convexe en x et en y tend pour n croissant vers $(\exp n^2(\lambda-\lambda^2) \bar{F}_\lambda)$ ou $\bar{F}_\lambda$ est le maximum en x,y de $F_\lambda(x,y)$.

D'après (8') et (9) le résultat sera donc établi quand on pourra prouver que :

(10) $$\bar{F}_\lambda < 2 \operatorname{Log} 2$$

Observons d'abord que $F(x,y)$ n'atteint son maximum que si $x = y$. Dans ce cas : $F_\lambda(x,x) = 2H(\frac{x\lambda}{1-\lambda}) + \frac{2\lambda}{1-\lambda}\, x\, L\, 2 + \frac{\lambda}{1-\lambda} H(2x)$ qui pour $x = 1/2$ est certainement plus grand que 2 log 2 quand λ est voisin de 1/2. Par contre, $F_\lambda(x,x)$ croissant en λ, tend uniformément vers zéro avec cette quantité. Il existe donc une valeur λ_0 finie telle que (2°) soit satisfaite à la limite quand $n \longrightarrow \infty$ si $\lambda < \lambda_0$.

Références sur les treillis

G. BIRKHOFF. Lattice Theory, New York 1948, Chap. III.

M. L. DUBREIL-JACOTIN, , L. LESIEUR et R. CROISOT -Theorie des Treillis (Paris 1953).