Can. J. Math., Vol. XXIX, No. 3, 1977, pp. 631–637

MATRICES SOUS-STOCHASTIQUES ET FONCTIONS CONVEXES

PAL FISCHER ET JOHN A. R. HOLBROOK

1. Introduction. Dans cette note on se propose d'étendre certains résultats de L. Mirsky concernant des matrices doublement sous-stochastiques [**5**; **6**; **7**].

On commence par rappeler des définitions et notations qu'on aura à utiliser par la suite.

Une matrice (carrée) M d'ordre n est dite *doublement sous-stochastique* (d.s.s.) si ses éléments m_{ij} vérifient les conditions suivantes:

(1) $\quad m_{ij} \geqq 0 \quad (i, j = 1, \ldots, n),$

(2) $\quad \sum_{j=1}^{n} m_{ij} \leqq 1 \quad (i = 1, \ldots, n),$ et

(3) $\quad \sum_{i=1}^{n} m_{ij} \leqq 1 \quad (j = 1, \ldots, n).$

Lorsqu'il y a égalité en (2) et (3), la matrice M est appelée *doublement stochastique* (d.s.). (Une telle matrice est connue aussi sous le nom d'une matrice *bistochastique*.)

Soit $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. On désigne par $x^* = (x_1^*, \ldots, x_n^*) \in \mathbf{R}^n$ le réarrangement décroissant de x, c'est-à-dire, $x_1^*, \ldots, x_n^*$ est une permutation de $x_1, \ldots, x_n$ telle que $x_1^* \geqq \ldots \geqq x_n^*$. Soient $x, y \in \mathbf{R}^n$. On écrit $x \gg y$ pour indiquer que

(4) $\quad \sum_{i=1}^{k} x_i^* \geqq \sum_{i=1}^{k} y_i^* \quad (k = 1, \ldots, n),$

et on écrit $x \succ y$ si $x \gg y$ et si l'on a l'égalité dans (4) pour $k = n$.

L'origine de ce travail se trouve dans le théorème suivant, dû à L. Mirsky [**5**; **6**, Theorem 3a; **7**; Section 6].

THÉORÈME 1. *Soient $x, y \in \mathbf{R}^n$. Les quatre conditions suivantes sont équivalentes*

(a) $x^+ \gg y$ *et* $(-x)^+ \gg (-y)$ *où* $x^+ = (x_1^+, \ldots, x_n^+)$ *et* $x_k^+ = \max(x_k, 0)$;

(b) *il existe une matrice d.s.s. M telle que* $y = Mx$;

(c) $y \in \operatorname{conv}\{(\delta_1 x_{\sigma(1)}, \ldots, \delta_n x_{\sigma(n)})$: *$\sigma$ est une permutation de l'ensemble* $(1, \ldots, n)$ *et* $\delta_k = 0$ *ou* $1\}$;

Reçu le 12 juillet, 1976. Ce travail a été subventionné partiellement par le C.N.R. du Canada (subvention A8745).

https://doi.org/10.4153/CJM-1977-064-5 Published online by Cambridge University Press

(d) *quelle que soit la fonction continue et convexe* $f : \mathbf{R} \to \mathbf{R}$ *telle que* $f \geqq f(0)$, *on a*

$$\sum_1^n f(x_k) \geqq \sum_1^n f(y_k).$$

Ici, on désigne par conv A *l'enveloppe convexe de l'ensemble* A.

À son tour, ce théorème fût inspiré, selon Mirsky, par un résultat classique de G. H. Hardy, J. E. Littlewood, G. Pólya, et R. Rado [**3**; **9**]. Ceci se peut présenter de la façon suivante.

THÉORÈME 2. *Soient* $x, y \in \mathbf{R}^n$. *Les conditions suivantes sont équivalentes:*
(a′) $x \succ y$;
(b′) *il existe une matrice d.s. telle que* $y = Mx$;
(c′) $y \in \operatorname{conv}\{(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) : \sigma$ *est une permutation*$\}$;
(d′) *quelle que soit la fonction continue et convexe* $f: \mathbf{R} \to \mathbf{R}$, *on a*

$$\sum_1^n f(x_k) \geqq \sum_1^n f(y_k). \tag{5}$$

Il y a une généralisation bien connue (voir, par exemple, Remarque 6) du théorème 2 à plusieurs variables, mais apparemment on ne connait aucune généralisation pareille du théorème 1. Le but de cet article est de combler cette lacune.

2. Résultats auxiliaires. Dans le reste du travail actuel on suppose que E est un espace linéaire réel ($\mathbf{R}^d$, par exemple). La proposition suivante vérifie que la relation (b) $\Rightarrow$ (d) se peut étendre au cas de plusieurs variables.

PROPOSITION 3. *Soient* x_k, y_k *des éléments de l'espace* E *tels qu'il y a une matrice d.s.s.* M *satisfaisant à la relation* $y = Mx$, *c'est-à-dire,*

$$y_i = \sum_{j=1}^n m_{ij} x_j \quad (i = 1, \ldots, n). \tag{6}$$

Soit f *une fonction convexe telle que* $f : E \to \mathbf{R}$ *et* $f \geqq f(0)$. *Alors* f *satisfait à l'inégalité* (5).

Démonstration. Il est suffisant de considérer le cas où $f(0) = 0$. En effet, dans le cas contraire on remplace f par $f - f(0)$ et on voit que f satisfait (5) si et seulement si $f - f(0)$ satisfait (5). Ainsi la convexité de f entraîne que

$$f(rz) = f(rz + (1 - r)0) \leqq rf(z) + (1 - r)f(0) = rf(z) \quad (z \in E) \tag{7}$$

pour $0 \leqq r \leqq 1$. Par conséquent, si l'on désigne par r_i la somme $\sum_{j=1}^n m_{ij}$ $(i = 1, \ldots, n)$, on a

$$\begin{aligned} f(y_i) &= f\left(r_i \sum_{j=1}^n \frac{m_{ij}}{r_i} x_j\right) \\ &\leqq r_i f\left(\sum_{j=1}^n \frac{m_{ij}}{r_i} x_j\right) \leqq r_i \sum_{j=1}^n \frac{m_{ij}}{r_i} f(x_j) = \sum_{j=1}^n m_{ij} f(x_j), \end{aligned}$$

dans le cas où $r_i > 0$. Comme la relation

$$f(y_i) \leqq \sum_{j=1}^{n} m_{ij} f(x_j)$$

est triviale dans le cas où $r_i = 0$, on voit que

$$\sum_1^n f(y_i) \leqq \sum_{i,j=1}^{n} m_{ij} f(x_j) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} m_{ij} \right) f(x_j) \leqq \sum_1^n f(x_j).$$

Dans la démonstration de notre théorème principal nous aurons besoin du résultat suivant sur le "balayage" correspondant à un cône convexe de fonctions. Apparemment ce résultat est dû à plusieurs auteurs, notamment à H. Bauer, P. Cartier, J. M. G. Fell, P. A. Meyer, et G. Mokobodzki; ici nous suivons la présentation de Meyer [**4**, chapitre XI].

THÉORÈME 4. *Soit X un espace compact et métrisable, et soit C un cône convexe de fonctions continues réelles défines sur X. On suppose que C contient la fonction constante* 1 *et que*

(8) $\quad f, g \in C \Rightarrow \min (f, g) \in C.$

Soient μ et λ deux mesures (non-négatives, boréliennes) sur X. Les conditions suivantes sont équivalentes.

(α) *Quelle que soit la fonction $f \in C$, on a*

$$\int f d\mu \leqq \int f d\lambda.$$

(β) *Pour chaque $x \in X$ on peut correspondre une mesure T_x sur X telle que d'une part*

(9) $\quad f(x) \geqq \int f dT_x \quad (f \in C),$

et d'autre part, pour toute fonction continue $g: X \rightarrow \mathbf{R}$, la relation $h(x) = \int g dT_x$ définie une fonction borélienne et, de plus, on a

(10) $$\int \left(\int g dT_x \right) d\lambda(x) \left(= \int h d\lambda \right) = \int g d\mu.$$

On aura besoin seulement de la conséquence suivante du théorème ci-dessus.

COROLLAIRE 5. *Soient x_k, y_k $(k = 1, \ldots, n)$ des éléments dans l'espace E. Soit C^* un cône convexe de fonctions réélles définies sur E tel que C^* contient la fonction constante* -1 *et*

$$f, g \in C^* \Rightarrow \max (f, g) \in C^*.$$

Si pour toute fonction $f \in C^*$ on a

(11) $\quad \sum_1^n f(x_k) \geqq \sum_1^n f(y_k),$

https://doi.org/10.4153/CJM-1977-064-5 Published online by Cambridge University Press

alors il existe une matrice d.s. M telle que

$$(12) \quad f(y_i) \leqq \sum_{j=1}^{n} m_{ij} f(x_j) \quad (i = 1, \ldots, n; f \in C^*).$$

Démonstration. Soit $X = \{x_k, y_k : k = 1, \ldots, n\}$ muni de la topologie discrète. De plus, soit

$$C = \{-f|X : f \in C^*\},$$

où $-f|X$ désigne la restriction de $-f$ à X. On voit que X et C vérifient les conditions du théorème 4. Soient

$$\mu = \sum_{k=1}^{n} \delta_{x_k}, \quad \lambda = \sum_{k=1}^{n} \delta_{y_k},$$

où l'on désigne par δ_z la mesure formée d'une masse $+1$ au point $z \in X$. Alors la condition (11) entraîne (α) du théorème 4, et d'après ce théorème, il y a des mesures T_x ayant les propriétés (9) et (10). Soient $z_1, \ldots, z_m$ les éléments distincts de l'ensemble $\{x_1, \ldots, x_n\}$. Avec $f \equiv 1$, la relation (9) entraîne

$$(13) \quad 1 \geqq \sum_{k=1}^{m} T_{y_i}(\{z_k\}) \quad (i = 1, \ldots, n).$$

D'autre part, si l'on prend g égale à la fonction indicatrice de $\{z_k\}$ on obtient de (10) la relation

$$(14) \quad \sum_{i=1}^{n} T_{y_i}(\{z_k\}) = \#\{j : x_j = z_k\}.$$

Ainsi, on a

$$(15) \quad \sum_{i=1}^{n} \left(\sum_{k=1}^{m} T_{y_i}(\{z_k\}) \right) = \sum_{k=1}^{m} \#\{j : x_j = z_k\} = n.$$

De (13) et (15) on conclut que

$$(16) \quad \sum_{k=1}^{m} T_{y_i}(\{z_k\}) = 1 \quad (i = 1, \ldots, n).$$

Il s'ensuit que le support de chaque mesure T_{y_i} est un sous-ensemble de l'ensemble $\{z_1, \ldots, z_m\}$ parce que (9) (avec $f \equiv 1$) montre que la mesure totale de T_x est au plus égale à 1. Donc

$$(17) \quad f(y_i) \geqq \sum_{k=1}^{m} T_{y_i}(\{z_k\}) f(z_k) \quad (i = 1, \ldots, n; f \in C).$$

Finalement, on pose

$$m_{ij} = \frac{T_{y_i}(\{x_j\})}{\#\{k : x_k = x_j\}}.$$

Les relations (14) et (16) entraînent que $M = [m_{ij}]$ est une matrice d.s., pendant que (12) découle de (17) en vertu de la relation entre C^* et C.

Remarque 6. Il est bien connu que résultats tels que le théorème 4 et son corollaire constituent généralisations de l'équivalence (b′) $\Leftrightarrow$ (d') du théorème 2. En effet, si E est un espace vectoriel topologique localement convexe et x, y sont n-tuples d'éléments $x_k, y_k \in E$, il s'ensuit facilement du corollaire 5 que les conditions suivantes sont équivalentes:

(B′) il existe une matrice M d.s. telle que $y = Mx$;
(D′) quelle que soit la fonction continue et convexe $f : E \rightarrow \mathbf{R}$ on a

$$\sum_1^n f(x_k) \geqq \sum_1^n f(y_k).$$

À l'aide d'un calcul élémentaire on peut vérifier que (B′) $\Rightarrow$ (D′). D'autre part, en posant

$$C^* = \text{les fonctions réelles continues et convexes sur } E,$$

on voit que la relation (12) résulte du corollaire 5. Puisque pour chaque forme linéaire continue x^* sur E, on a x^*, $-x^* \in C^*$, (12) entraîne

$$x^*(y_i) = \sum_{j=1}^n m_{ij}x^*(x_j) = x^*\left(\sum_{j=1}^n m_{ij}x_j\right) \quad (i = 1, \ldots, n).$$

Donc, par le théorème de la séparation,

$$y_i = \sum_{j=1}^n m_{ij}x_j \quad (i = 1, \ldots, n),$$

c'est-à-dire, on a (B′).

Remarque 7. L. Mirsky a démontré dans [**5**] que, si $x, y \in \mathbf{R}^n$, pour que soit $x \gg y$ il faut et il suffit que l'on ait une matrice M d.s. telle que $y \leqq Mx$, c'est-à-dire,

$$(18) \quad y_i \leqq \sum_{j=1}^n m_{ij}x_j \quad (i = 1, \ldots, n).$$

Bien que cela découle facilement du théorème 2, il est intéressant de préciser la relation qui existe entre ce résultat, le corollaire 5, et un théorème de G. Pólya [**8**] (voir aussi K. M. Chong [**2**, Theorem 2.6]). Selon Pólya, $x \gg y$ entraîne la condition (11) du corollaire 5 avec

$$C^* = \text{les fonctions continues, convexes et non-décroissantes sur } R.$$

Alors, il ne faut que prendre $f(t) = t$ dans les relations (12) pour déduire (18).

3. Théorème principal. Nous sommes maintenant en mesure de prouver l'extension du théorème 1 pour plusieurs variables.

THÉORÈME 8. *Soit E un espace vectorial topologique localement convexe et soient x, y deux n-tuples d'éléments x_k, $y_k \in E$ ($k = 1, \ldots, n$). Les conditions suivantes sont équivalentes:*

(B) *il existe une matrice M d.s.s. telle que $y = Mx$;*

(C) $y \in \operatorname{conv}\{(\delta_1 x_{\sigma(1)}, \ldots, \delta_n x_{\sigma(n)}) : \sigma$ est une permutation et $\delta_k = 0$ ou $1\}$;

(D) *quelle que soit la fonction continue et convexe $f : E \to \mathbf{R}$ telle que $f \geqq f(0)$, on a*

$$\sum_{k=1}^{n} f(x_k) \geqq \sum_{k=1}^{n} f(y_k).$$

Démonstration. (C) $\Rightarrow$ (B). La matrice S, définie par

$$s_{ij} = \begin{cases} \delta_i & \text{pour } j = \sigma(i) \\ 0 & \text{pour } j \neq \sigma(i), \end{cases}$$

est une matrice d.s.s., et elle fait correspondre à x le n-tuple

$$(19) \quad (\delta_1 x_{\sigma(1)}, \ldots, \delta_n x_{\sigma(n)}).$$

Donc, si y est une combinaison convexe des vecteurs du type (19), il suffit de définir M comme la même combinaison convexe des matrices S.

(B) $\Rightarrow$ (D). Cela résulte de la proposition 3.

(D) $\Rightarrow$ (C). D'après le théorème de la séparation il suffit de prouver que, pour chaque forme linéaire continue ϕ sur E^n, il existe un n-tuple ν du type (19) tel que $\phi(y) \leqq \phi(\nu)$. Puisque chaque ϕ a la forme $\phi(\nu) = \sum_{i=1}^{n} \phi_i(\nu_i)$ où les ϕ_i sont des formes linéaires continues sur E, il est suffisant de montrer que l'on peut trouver une permutation σ et $\delta_i (=0$ ou $1)$ tels que l'on ait

$$(20) \quad \sum_{i=1}^{n} \phi_i(y_i) \leqq \sum_{i=1}^{n} \phi_i(\delta_i x_{\sigma(i)}).$$

Pour cela on applique le corollaire 5 avec

$$C^* = \text{les fonctions continues et convexes } f : E \to \mathbf{R}, \text{ telles que } f \geqq f(0).$$

D'après ce corollaire (D) entraîne (11) pour toute fonction $f \in C^*$. Comme $\phi_i^+ (= \max(\phi_i, 0)) \in C^*$, il découle que

$$(21) \quad \sum_{i=1}^{n} \phi_i(y_i) \leqq \sum_{i=1}^{n} \phi_i^+(y_i) \leqq \sum_{i=1}^{n} \left(\sum_{j=1}^{n} m_{ij} \phi_i^+(x_j) \right) = \sum_{i,j=1}^{n} m_{ij} \phi_i^+(x_j)$$

où $M = [m_{ij}]$ est une matrice d.s. qui vérifie les conditions du corollaire 5. En vertu d'un résultat classique de G. Birkhoff [**1**], M est une combinaison convexe de forme

$$\sum_{\sigma} \lambda_\sigma P^\sigma$$

https://doi.org/10.4153/CJM-1977-064-5 Published online by Cambridge University Press

où $P^\sigma = [P_{ij}{}^\sigma]$ est la matrice de permutation associée à la permutation σ par les relations

$$P_{ij}{}^\sigma = \begin{cases} 1 & \text{pour } j = \sigma(i) \\ 0 & \text{pour } j \neq \sigma(i). \end{cases}$$

Alors de (21) on déduit

$$(22) \quad \sum_{i=1}^{n} \phi_i(y_i) \leqq \sum_{\sigma} \lambda_\sigma \left(\sum_{i,j=1}^{n} P_{ij}{}^\sigma \phi_i{}^+(x_j) \right).$$

Donc il existe une permutation σ telle que

$$\sum_{i=1}^{n} \phi_i(y_i) \leqq \sum_{i,j=1}^{n} P_{ij}{}^\sigma \phi_i{}^+(x_j) = \sum_{i=1}^{n} \phi_i{}^+(x_{\sigma(i)}).$$

En posant

$$\delta_i = \begin{cases} 1 & \text{pour } \phi_i(x_{\sigma(i)}) > 0 \\ 0 & \text{pour } \phi_i(x_{\sigma(i)}) \leqq 0. \end{cases}$$

on obtient la relation (20).

Bibliographie

1. G. Birkhoff, *Tres observaciones sobre el álgebra lineal*, Rev. Univ. Nac. Tucumán *A5* (1946), 147–151.
2. K. M. Chong, *An induction theorem for rearrangements*, Can. J. Math. *28* (1976), pp. 154–160.
3. G. H. Hardy, J. E. Littlewood, et G. Pólya, *Inequalities*, second edition (Cambridge U.P., 1952).
4. P. A. Meyer, *Probability and potentials* (Blaisdell Publishing Company, 1966).
5. L. Mirsky, *On a convex set of matrices*, Archiv der Math. *10* (1959), 88–92.
6. ——— *Inequalities for certain classes of convex functions*, Proc. Edinburgh Math. Soc. *11* (1958/9), 231–235.
7. ——— *Majorization of vectors and inequalities for convex functions*, Monatshefte für Mathematik *65* (1961), 159–169.
8. G. Pólya, *Remark on Weyl's note: Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Nat. Acad. Sci. *36* (1950), 49–51.
9. R. Radó, *An inequality*, J. London Math. Soc. *27* (1952), 1–6.

Université de Guelph,
Guelph, Ontario

https://doi.org/10.4153/CJM-1977-064-5 Published online by Cambridge University Press