REVUE FRANÇAISE D'AUTOMATIQUE, D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE. RECHERCHE OPÉRATIONNELLE

JULIO DIAZ
JAIME GONZÀLEZ
NANCY LACOURLY

Construction de rankings dans un tournoi

Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle, tome 23, nº 1 (1989), p. 51-56.

<http://www.numdam.org/item?id=RO_1989__23_1_51_0>

Recherche opérationnelle/Operations Research
(vol. 23, n° 1, 1989, p. 51 à 56)

CONSTRUCTION DE RANKINGS DANS UN TOURNOI (*)

par Julio DIAZ [1], Jaime GONZÀLEZ [1] et Nancy LACOURLY [1]

Résumé. — *Nous proposons une méthode métrique pour construire des rankings pour n joueurs jouant dans un tournoi incomplet et non balancé. La méthode consiste à ajuster les résultats du tournoi de manière à produire une relation transitive sur les n joueurs et ainsi obtenir un ranking. Le critère proposé nous conduit à résoudre un problème de programmation non linéaire à* $n-1$ *variables bornées.*

Mots clés : Ranking; tournoi; incomplet; non balancé.

Abstract. — *We propose a metric method to rank n players in an incomplete, unbalanced tournament. The method adjusts the tournament outcomes, in order to produce a transitive relation, and a ranking for the n players. The proposed criterion requires solving a non-linear programming problem on* $n-1$ *bounded variables.*

Keywords : Ranking; tournament; incomplete; unbalanced.

1. INTRODUCTION

Pour classifier les joueurs qui s'affrontent dans un tournoi de nombreuses méthodes ont été proposées, entre autres les métriques, qui fournissent une évaluation numérique à chaque joueur, appelée ranking; celle-ci permet de mesurer les différences entre joueurs (par exemples, les méthodes de Bradley-Terry [2], Jech [3] et Thurstone [5]. Cette dernière ne s'applique pas aux tournois incomplets et les deux premières attribuent le même ranking aux joueurs qui obtiennent le même score.

(*) Reçu septembre 1988
Travail financé par FONDECYT, Projet N° 353.
(1) Département de Mathémathiques, Université du Chili, Casilla 170/3, Santiago, Chili.

2. FORMULATION DU PROBLÈME

Considérons les résultats d'un tournoi entre n joueurs ou équipes $E_1, E_2, \ldots, E_n$. Après une partie, les joueurs E_i et E_j obtiennent respectivement les points t_i et t_j, tels que $t_i=1$ et $t_j=0$, si E_i bat E_j et $t_i=t_j=1/2$, si E_i y E_j égalisent.

Si m_{ij} est le nombre de parties jouées entre E_i et E_j et r_{ij} est la somme des points obtenus par E_i au cours de ces m_{ij} parties contre E_j, la matrice $R=(r_{ij})$ des résultats vérifie

$$\begin{gathered} r_{ij} \geqq 0 \text{ pour } i, j=1, \ldots, n \\ r_{ij}+r_{ji}=m_{ij} \text{ pour } i, j=1, \ldots, n. \\ r_{ii}=0, \text{ pour } i=1, \ldots, n. \end{gathered} \tag{1}$$

Quand $m_{ij} \neq 0$, pour tout $i \neq j$, le tournoi est *complet*, et en particulier si $m_{ij}=m$, on parle de tournoi *balancé*.

Le *score* du joueur E_i est la somme de tous ses points

$$s_i = \sum_j r_{ij} \tag{2}$$

Plus le score est élevé, meilleur est le joueur. On pourrait donc envisager de prendre les scores comme rankings pour indiquer les forces relatives des joueurs. Mais les scores ne respectent pas toujours le principe de Condorcet, qui postule que si les résultats sont transitifs, alors les rankings doivent suivre le même ordre.

Si $u=(u_i)$ est un vecteur de rankings, on doit avoir: « le joueur i est meilleur que le joueur j si et seulement si $u_i > u_j$ ». La relation « meilleur que » est donc transitive.

Si P_{ij} est la probabilité que le joueur i gagne le joueur j

$$(0 < P_{ij} < 1,\ P_{ij} + P_{ji} = 1),$$

alors le joueur i est « meilleur que » le joueur j si et seulement si $P_{ij} > 1/2$.

Il faut alors imposer la condition de transitivité: si $P_{ij} > 1/2$ et

$$P_{jk} > 1/2 \Rightarrow P_{ik} > 1/2.$$

A partir des résultats r_{ij}, il nous faut donc déterminer les probabilités P_{ij} et les rankings u_i associés.

Un estimateur raisonnable de P_{ij} serait :

$$\hat{P}_{ij}=r_{ij}/(r_{ij}+r_{ji})\,;$$

Malheureusement, ces estimateurs violent généralement la condition de transitivité et de plus ne permettent pas d'estimer tous les P_{ij} dans le cas de tournois incomplets.

Il y a une manière naturelle de lier les rankings u_i et les probabilités P_{ij} :

$$P_{ij}=u_i/(u_i+u_j) \qquad \text{avec} \quad u_i>0\ . \tag{3}$$

Dans ce cas, les p_{ij} vérifient les conditions des probabilités et la condition de transitivité. Jech [3] donne une condition nécessaire et suffisante pour qu'il existe un vecteur u qui vérifie (3) :

$$P_{ij}P_{jk}P_{ki}=P_{ji}P_{ik}P_{kj}, \qquad i,\, j,\, k\in\{1,\ \ldots,\ n\},\ i\neq j\neq k. \tag{4}$$

3. LA MÉTHODE PROPOSÉE

Les méthodes de Thurstone Modifié [1], Bradley-Terry [2], construisent des rankings en ajustant le mieux possible les scores s_i,

$$\sum_j m_{ij}P_{ij}=s_i, \qquad \text{pour tout} \quad i=1,\ \ldots,\ n.$$

Les rankings dépendent alors du modèle probabilistique (Normal pour Thurstone et Logistique pour Bradley). Jech ajuste aussi les scores avec les conditions (3) et (4) et obtient les mêmes rankings que Bradley.

Notre méthode, comme celle de Thurstone, estime les P_{ij} à partir des résultats r_{ij} avec la condition (3) de Jech :

$$\mathrm{Min}\sum_i\sum_j(r_{ij}-m_{ij}P_{ij})^2 \tag{5}$$

avec

$$P_{ij}P_{jk}P_{ki}=P_{ji}P_{ik}P_{kj}, \qquad i,\, j,\, k=1,\ \ldots,\ n, \quad i\neq j\neq k,$$
$$P_{ij}+P_{ji}=1,$$
$$0\leqq P_{ij}\leqq 1.$$

THÉORÈME 1 : *La matrice P est complètement déterminée si on connait une ligne de P dont les composantes sont différentes de* 0 *et* 1.

Démonstration: Supposons que la ligne k de P contienne des éléments différents de 0 et 1. Alors la condition (5) implique

$$P_{ij} P_{jk} P_{ki} = P_{ik} P_{kj} P_{ji} \qquad \text{pour tout} \quad i \neq j \neq k$$

et comme $P_{rs} + P_{sr} = 1$ pour tout $r \neq s$, on en déduit que

$$P_{ij} P_{ki} (1 - P_{kj}) = (1 - P_{ij}) P_{kj} (1 - P_{ki})$$

et donc pour $i \neq j \neq k$

$$P_{ij} = P_{kj} (1 - P_{ki}) / (P_{ki} (1 - P_{kj}) + P_{kj} (1 - P_{ki})) \tag{6}$$

ce qui montre que la ligne i est totalement définie en fonction de la ligne k. □

Notons que si le quotient de (6) est non nul, les p_{ij} sont bornées entre 0 et 1; il suffit donc d'utiliser des critères d'arrêt pas trop exigents.

THÉORÈME 2 : *Le problème* (5) *équivaut à minimiser une fonction à* $n-1$ *variables sous des restrictions linéaires.*

Démonstration: Observons que:

$$r_{kj} - m_{kj} P_{kj} = (m_{jk} - r_{jk}) - m_{jk} (1 - P_{jk}) = m_{jk} P_{jk} - r_{jk} \tag{7}$$

et comme $m_{ii} = 0$ et $r_{ii} = 0$, on a

$$\begin{aligned}\sum_i \sum_j (r_{ij} - m_{ij} P_{ij})^2 &= \sum_{i>j} \sum_j 2 (r_{ij} - m_{ij} P_{ij})^2 \\ &= \sum_{i>j} \sum_j 2 (r_{ij} - m_{ij} P_{ij} (P_{ki}, P_{kj}))^2 \\ &= f(P_{k1}, P_{k2}, \ldots, P_{kn})^2\end{aligned}$$

où $P_{ij}(P_{ki}, P_{kj})$ est donnée par (6).

Comme les restrictions d'égalité du problème (8) ont été incorporées explicitement dans la fonction objective à partir des expressions (6) et (7), le problème (5) équivaut à

$$\operatorname{Min} f(P_{k1}, P_{k2}, \ldots, P_{kn}) \tag{8}$$

avec $0 \leqq P_{ki} \leqq 1$ pour $i = 1, \ldots, n$, $i \neq k$ ce qui démontre la proposition. □

Le problème (8) est un problème de Programmation non linéaire dont les variables sont bornées et pour le résoudre on a modifié l'algorithme général décrit dans [4]. Il est facile de voir qu'il existe toujours une solution, f étant continue et l'ensemble des restrictions un compacte non vide.

On déduit finalement un vector ranking u^* de la solution $P^*=(P_{ij}^*)$ de (8)

$$u_i^*=(P_{ij}^*/P_{ji}^*)\,u_j^* \qquad \text{avec} \quad u_1^*+u_2^*+\ldots+u_n^*=K \qquad (\text{e. g., } K=100)$$

4. EXEMPLES

Nous présentons deux exemples et nous comparons les rankings obtenus avec ceux de Jech.

Les résultats du tableau I proviennent d'un tournoi incomplet balancé et ceux du tableau III, d'un tournoi incomplet non balancé. Les rankings obtenus dans le premier cas sont semblables à ceux Jech (tableau II), alors que

TABLEAU I

Résultats du tournoi				Score obtenu
0,0	2,0	2,0	0,0	4,0
1,0	0,0	0,0	3,0	4,0
1,0	0,0	0,0	2,0	3,0
0,0	0,0	1,0	0,0	1,0

TABLEAU II

Matrice de probabilités				Score espéré	Ranking	Ranking de Jech
0,0	0,60	0,76	0,90	4,06	47,54	46,25
0,40	0,00	0,67	0,86	3,77	31,72	35,79
0,26	0,33	0,00	0,74	2,97	15,44	13,85
0,10	0,14	0,26	0,00	1,20	5,30	4,12

TABLEAU III

Résultats du tournoi				Score obtenu
0,0	1,0	0,0	0,5	1,5
0,0	0,0	1,0	1,0	2,0
0,0	0,0	0,0	1,0	1,0
0,5	0,0	0,0	0,0	0,5

TABLEAU IV

Matrice de probabilités				Score espéré	Ranking	Ranking de Jech
0,0	0,76	0,97	1,00	1,76	74,09	48,83
0,24	0,00	0,91	1,00	2,15	23,55	33,40
0,03	0,09	0,00	0,99	1,08	2,33	12,78
0,00	0,00	0,01	0,00	0,01	0,03	0,14

dans le deuxième cas (tableau IV), ils sont assez différents. Ceci est dû au fait que les résultats du tableau I sont transitifs et ceux du tableau III ne le sont pas. Dans ce cas notre méthode pénalise le dernier joueur qui n'a jamais gagné et récompense le premier qui n'a jamais perdu.

BIBLIOGRAPHIE

1. M. BACCELIERE, *Clasificación con datos de preferencias*, Mémoire d'Ingénieur Civil Mathématique, Université du Chili, 1981.
2. R. BRADLEY et M. TERRY, *Rank Analysis of Incomplete Block Design, The Method of paired Comparisons*, Biometrika, vol. 39, 1952, pp. 324-345.
3. T. JECH, *The Ranking of Incomplete Tournaments: a Mathematicians Guide to Popular sports*, A.M.S., 1983, pp. 246-266.
4. G. MC CORMICK, *Non Linear Programming Theory, Algorithms and Applications*, Wiley, 1983.
5. L. THURSTONE, *The Prediction of Choice*, Psychometrika, vol. 10, 1945, pp. 237-253.