SUR LA THÉORIE DE LA RECHERCHE ALÉATOIRE

par Monsieur le Professeur A. RENYI

Académie des Sciences de Hongrie

RESUME

Nous désirons trouver un élément inconnu x, d'un ensemble fini S en observant certaines fonctions f_1, f_2,..., au point x. Les fonctions f_i sont choisies appartenant à un ensemble F. Dans cet article, il est montré qu'une "stratégie mixte" qui consiste à choisir les fonctions f_i de F d'une manière aléatoire peut être, sous certaines conditions, presque aussi bonne que la meilleure "stratégie pure".

SUMMARY

This paper deals with the following problem : we want to find an unknown element x of a finite set S by observing certain functions $f_1(x)$, $f_2(x)$,... at the point x. The functions are chosen from a set F of functions ; in the paper it is shown that a "mixed strategy" consisting in choosing the functions f_i from F at random may be under certain conditions almost as good as the best pure strategy.

INTRODUCTION

Les problèmes de recherche, traités dans cet article, peuvent être décrits par le simple exemple suivant :

Soit S_n un ensemble fini ayant $n \geqslant 2$ éléments distincts -appelés points- et supposons que nous voulions trouver un point inconnu x de l'ensemble S_n. L'ensemble S_n lui-même nous est supposé connu ; Supposons en outre qu'il soit impossible d'observer directement x, mais que cependant nous puissions choisir des fonctions f_1, f_2,..., f_N appartenant à un ensemble donné F de fonctions définies sur S_n, et observer les valeurs $f_1(x)$, $f_2(x)$,..., $f_N(x)$ prises par ces fonctions pour l'inconnu x. Bien entendu, si F pouvait contenir une fonction f qui prendrait des valeurs toutes différentes en des points différents, une unique observation de cette fonction serait suffisante. Nous supposons cependant que toute fonction f appartenant à la classe F est telle que le nombre de valeurs différentes prises par f est très inférieur à n. (Nous serons spécialement intéressés par le cas où $f \in F$ ne prend que les valeurs 0 et 1 et où n est un grand nombre). Dans un tel cas, naturellement, il est nécessaire d'observer les valeurs d'un grand nombre de fonctions f au point x.

Chacune de ces observations ne nous donne qu'une information partielle sur x (en fait elle détermine un sous-ensemble A de S_n dont x doit faire partie) ; mais après avoir fait un nombre assez grand de telles observations, l'information obtenue s'accumule et nous permet de déterminer x par un nombre d'observations qui ne soit pas trop grand. Nous pouvons aussi supposer qu'à chaque observation correspond une certaine dépense (ou qu'elle demande un laps de temps défini) et désirer limiter la dépense (ou la durée) de toute la procédure de recherche.Nous appellerons stratégie de recherche une méthode de choix successif des fonctions $f_1,\ldots, f_N$, qui conduit, à son terme, à la détermination de l'inconnu x. Evidemment, on essaye en pratique de choisir une stratégie avec N (nombre de fonctions à observer) aussi petit que possible. De deux procédures de recherche, celle de moindre durée (en moyenne) est la meilleure, cependant on peut avoir d'autres exigences. Par exemple une stratégie simple qui peut également être aisément programmée sur ordinateur est souvent préférable à une stratégie compliquée. Si A et B sont deux stratégies telles que A nécessite (en moyenne) l'observation d'un plus petit nombre de fonctions que B (A "meilleur" que B) mais que, par contre, l'exécution pratique de A soit beaucoup plus longue et pénible, alors B pourra, pour des raisons pratiques, être préféré à A.

Une stratégie sera appelée "stratégie pure" si elle fixe de façon unique le choix des fonctions $f_1, f_2,\ldots, f_N$, et "stratégie mixte" si le choix de ces fonctions est aléatoire. Une stratégie pure sera appelée "prédéterminée" si le nombre N et le choix de chacune des fonctions $f_1, f_2,\ldots, f_N$ sont déterminés par avance, avant de commencer les observations. Elle sera appelée "séquentielle" si seul le choix de f_1 est décidé par avance, le choix de f_k $(k \geqslant 2)$ étant fait seulement après avoir observé $f_1(x), f_2(x),\ldots, f_{k-1}(x)$ et pouvant dépendre des valeurs observées ; dans le cas d'une stratégie séquentielle (pure) le nombre des observations N dépend aussi, ordinairement, de la valeur de x.

Les problèmes de recherche touchent en pratique chaque domaine de l'activité humaine. Ses exemples typiques sont : diagnostics médicaux, analyses chimiques, recherche d'un défaut dans un mécanisme compliqué, recherche d'un objet perdu ou caché, recherche d'une faute dans une longue série d'évaluations ou dans un programme pour ordinateur, recherche de certaines données bibliographiques, recherche des racines d'une équation, du maximum d'une fonction, d'un paramètre, d'une distribution de probabilité, etc.

La théorie de la recherche devrait être considérée -du point de vue de l'auteur- comme un chapitre de la théorie de l'information ; ce chapitre n'est toutefois pas développé très profondément. Beaucoup d'intéressants problèmes particuliers de recherche aléatoire ont été examinés, mais une étude systématique de tels problèmes n'a pas encore été réalisée.

1. SYSTEMES SEPARANTS DE FONCTIONS

Evidemment pour qu'il puisse exister au moins une stratégie qui conduise à trouver l'élément inconnu x de S_n quel qu'il soit, il est nécessaire que la classe F des fonctions f disponibles soit assez riche.

Introduisons la définition suivante :

(I) - On dit qu'un système F de fonctions définies sur l'ensemble S_n sépare les éléments de S_n (ou, pour abréger, est un système séparant) s'il a la propriété suivante :

Pour chaque couple d'éléments différents x_1, x_2, $(x_2 \neq x_1)$ de S_n, il existe dans F une fonction f telle que $f(x_2) \neq f(x_1)$.

LEMME 1.– Soit F un système séparant minimal de fonctions qui sépare les éléments de l'ensemble fini à n éléments S_n. Si m est le nombre d'éléments de F, nous avons $m \leqslant n - 1$.

LEMME 2.– Si F est un système de fonctions qui sépare les éléments de l'ensemble S_n alors :

$$\sum_{f \in F} H(f) \geqslant \log n .$$

où H(f) est l'entropie de f.

2. DIFFERENTES NOTIONS D'HOMOGENEITE D'UN SYSTEME SEPARANT DE FONCTIONS

Dans nos investigations, nous ne supposons pas l'existence d'une quelconque structure algébrique, géométrique ou topologique sur l'ensemble S_n, et nous supposons de plus que nous n'avons pas de connaissance a priori de l'élément inconnu x de S_n, c'est-à-dire qu'avant de commencer la recherche toutes les possibilités $x = a_k$ $(k = 1, 2, \dots, n)$ ont la même probabilité a priori $1/n$. En conséquence, il est naturel de nous restreindre au cas dans lequel le système séparant F des fonctions disponibles f est, dans un certain sens, symétrique par rapport aux éléments a_k de S_n. Nous définirons différentes sortes de symétries de systèmes de fonctions. Tout d'abord, nous adoptons la définition suivante :

(II) - Un système F de fonctions f, définies sur l'ensemble :

$$S_n = \{a_1, a_2, \dots, a_n\} ,$$

est appelé complètement homogène, si notant par π chaque permutation de la suite 1, 2,... n pour chaque $f \in F$, la fonction g définie par :

$$g(a_j) = f(a_{\pi(j)})$$

appartient à F aussi.

Soit maintenant la définition suivante :

(III) - Un système de fonctions F défini sur l'ensemble $S_n = \{a_1, a_2, \dots, a_n\}$ est faiblement homogène d'ordre k ($2 \leqslant k \leqslant n$) s'il a la propriété suivante : Soient K éléments distincts $a_{j_1}, a_{j_2}, \dots, a_{j_k}$ de S_n et soit R_k le nombre de fonctions $f \in F$ telles que :

$$f(a_{j_1}) = f(a_{j_2}) = \dots = f(a_{j_K})$$

Alors R_k ne dépend pas du choix des éléments $a_{j1}, \dots, a_{jk}$.

Nous avons besoins aussi d'une autre propriété d'homogénéité qui est plus forte que l'homogénéité d'ordre k ; nous l'appelons forte homogénéité d'ordre k et nous la définissons ainsi :

(IV) - Soit F un ensemble de fonctions f définies sur l'ensemble $S_n = \{a_1, a_2, \dots, a_n\}$. Nous supposons que les valeurs de chaque $f \in F$ appartiennent à un ensemble $Y = \{y_1, \dots, y_r\}$. Nous appelons l'ensemble F fortement homogène d'ordre K si, choisissant K points distincts quelconques $(a_{j_1}, a_{j_2}, \dots, a_{j_K})$ de S_n et une suite quelconque d'éléments $y_{e_1}, \dots, y_{e_K}$ de Y [les y_{e_i} ($i = 1, 2, \dots, K$) ne sont pas nécessairement distincts], le nombre $R_K(y_{e_1} \dots y_{e_K})$ de ces fonctions $f \in F$ pour lesquelles :

$$f(a_{j_i}) = y_{e_i} \qquad (i = 1, 2, \dots, K)$$

ne dépend pas du choix des points $a_{j_1}, \dots, a_{j_K}$, mais il peut dépendre des valeurs $y_{e_1}, \dots, y_{e_K}$.

Nous introduisons alors une définition de plus :

(V) - Un système de fonctions F est dit homogène d'ordre K si, pour chaque décomposition $K = K_1 + K_2 + \dots + K_s$ de K, le nombre :

$$R\,[K_1, K_2, \dots, K_s]$$

de ces $f \in F$ pour lesquelles :

$$f(x_{11}) = f(x_{12}) = \dots\dots\dots\dots = f(x_{1K_1}),$$

$$f(x_{21}) = f(x_{22}) = \dots\dots\dots\dots = f(x_{2K_2}) \,,$$

$$f(x_{s1}) = f(x_{s2}) = \dots\dots\dots\dots = f(x_{sK_s}) \,,$$

ne dépend pas du choix des points x_{ij} de S_n s'ils sont tous distincts.

Maintenant nous prouvons le lemme suivant simple mais assez surprenant :

LEMME 3a.– L'ensemble F de fonctions f ne prenant que les valeurs 0 et 1, qui est faiblement homogène d'ordre 2, est aussi faiblement homogène d'ordre 3.

Il est aisé de prouver qu'un système faiblement homogène d'ordre k n'est pas nécessairement d'ordre $k + 1$ si $k \geqslant 3$. Ainsi l'affirmation du lemme 3a ne peut être généralisée en remplaçant 3 et 2 par $k + 1$ et k. Pourtant la généralisation suivante du lemme 3a est valide.

LEMME 3b.– Si le système F des fonctions prenant les valeurs 0 et 1 est faiblement homogène pour tout ordre $l \leqslant 2k$, alors il est aussi faiblement homogène d'ordre $2k + 1$ ($k = 1, 2, \dots$).

LEMME 4.– Soit F un système de fonctions prenant les valeurs 0 et 1 sur l'ensemble S_n, F étant faiblement homogène d'ordre 2, soit R_1 le nombre d'éléments de F et R_2 le nombre de fonctions $f \in F$ pour lesquelles $f(a_i) = f(a_j)$ $(i \neq j)$, alors nous avons :

$$\frac{R_2}{R_1} \geqslant \frac{(n-2)}{2(n-1)} \,.$$

THEOREME 1.– Si le système F de fonctions définies sur l'ensemble $S_n = (1, 2, \dots, n)$ est faiblement homogène d'ordre 2, nous avons pour tout $x \in S_n$:

$$P_1(n, N, F, x) \geqslant 1 - (n-1)\left(\frac{R_2}{R_1}\right)^N \,,$$

$$\text{et } P_2(n, N, F) \geqq 1 - \binom{n}{2}\left(\frac{R_2}{R_1}\right)^N$$

avec $P_1(n,N,F,x)$: probabilité que la suite $f_1(x), \dots, f_N(x)$ détermine x de façon unique, $P_2(n,N,F)$: probabilité que la suite $f_1(x), \dots, f_N(x)$ détermine x quel qu'il soit.

COROLLAIRE.– Si F_n est, pour tout n, un système séparant pour l'ensemble S_n faiblement homogène d'ordre 2, ayant $R_1(n)$ éléments et le paramètre $R_2(n)$ et si :

$$\lim_{n\to\infty}\left(N_n \log \frac{R_1(n)}{R_2(n)} - \log n\right) = +\infty ,$$

alors :

$$\lim_{n\to\infty} \min_{x\in S_n} P_1(n, N_n, F_n, x) = 1 .$$

Si nous supposons que :

$$\lim_{n\to\infty}\left(N_n \log \frac{R_1(n)}{R_2(n)} - 2 \log n\right) = +\infty ,$$

nous avons :

$$\lim_{n\to\infty} P_2(n, N_n, F_n) = 1 .$$

THEOREME 2a.– Si le système F des fonctions à valeurs 0 et 1 définies sur l'ensemble est faiblement homogène d'ordre 2 (et ainsi -lemme 3a- d'ordre 3) nous avons pour tout $x \in S_n$:

$$P_1(n, N, F, x) \leqslant 1 - (n-1)\left(\frac{R_2}{R_1}\right)^N + \binom{n-1}{2}\left(\frac{R_3}{R_1}\right)^N$$

COROLLAIRE.– Si, pour chaque n, F_n est un système de fonctions à valeurs 0 et 1 sur l'ensemble S_n faiblement homogène d'ordre 2,

si $R_2(n)/R_1(n) \leqslant 1/2$

et :

$$\lim_{n_j\to\infty}\left(N_{n_j} \log \frac{R_1(n_j)}{R_2(n_j)} - \log n_j\right) = C > 0 ,$$

alors, si x_{n_j} est un élément quelconque de S_{n_j},

$$1 - e^{-C} \leqslant \lim_{n_j\to\infty} \inf P_1(n_j, N_{n_j}, F_{n_j}, x_{n_j}) \leqslant \lim_{n\to\infty} \sup P_1(n_j, N_{n_j}, F_{n_j}, x_{n_j})$$

$$\leqslant 1 - e^{-C} + \frac{e^{-2C}}{2} ,$$

THEOREME 2b.– Si le système F est homogène d'ordre 4 sur l'ensemble S_n, alors nous avons :

$$P_2(n, N, F) \leqslant 1 - \binom{n}{2}\left(\frac{R_2}{R_1}\right)^N + \frac{3}{2}\binom{n}{3}\left(\frac{R_3}{R_1}\right)^N + 3\binom{n}{4}\left(\frac{R_4(2,2)}{R_1}\right)^N$$

THEOREME 3.— Si le système F est faiblement homogène pour chaque ordre, alors $P_1(n, N, F, x) = P_1(n, N, F)$ ne dépend pas de x et nous avons :

$$P_1(n, N, F) = \sum_{K=0}^{n-1} (-1)^k \binom{n-1}{K} \left(\frac{R_{K+1}}{R_1}\right)^N .$$

La preuve des résultats précédents ont été donnés dans l'article de l'auteur "Sur la théorie de la recherche aléatoire". (Bull. Amer. Math. Soc. 71 (1965) 809-828).