



**HAL**  
open science

## Sur les chaines à liaisons complètes linéaires

G. Mihôc

► **To cite this version:**

G. Mihôc. Sur les chaines à liaisons complètes linéaires. Annales de l'ISUP, 1961, X (3), pp.229-237. <hal-04092445>

**HAL Id: hal-04092445**

**<https://hal.science/hal-04092445v1>**

Submitted on 9 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# SUR LES CHAINES A LIAISONS COMPLÈTES LINÉAIRES

G. MIHÖC

I. Soit une suite de variables aléatoires

$$\xi_1, \xi_2, \dots, \xi_n, \dots$$

chacune pouvant prendre le même ensemble fini de valeurs  $\Omega = \{\omega_1, \omega_2, \dots, \omega_{m+1}\}$ , que nous allons supposer être des nombres réels.

On suppose en outre que l'on donne la répartition initiale

$$p_j = P(\xi_1 = \omega_j), \quad 1 \leq j \leq m+1.$$

Alors à la variable  $\xi_1$  correspond le schéma suivant :

$$\xi_1 : \begin{pmatrix} \omega_j \\ p_j \end{pmatrix} \quad 1 \leq j \leq m+1.$$

Notons avec

$$p_j^{(n)} = P(\xi_n = \omega_j); \quad (1 \leq j \leq m+1)$$

les probabilités absolues correspondant à la n-ième variable  $\xi_n$ .

On entend par probabilité absolue la probabilité que  $\xi_n$  prenne la valeur  $\omega_j$  lorsqu'on connaît toutes les valeurs prises par les variables précédentes

$$\xi_n : \begin{pmatrix} \omega_j \\ p_j^{(n)} \end{pmatrix} \quad 1 \leq j \leq m+1$$

$$p_{m+1}^{(n)} = 1 - \sum_{j=1}^m p_j^{(n)}$$

Par définition la suite de variables aléatoires  $\xi_n$  constitue une chaîne à liaisons complètes simple si les probabilités  $p_j^{(n+1)}$ ,  $1 \leq j \leq m+1$ , correspondant à la  $(n+1)$ -ième variable, ont la forme :

$$p_j^{(n+1)} = \psi_{ij}^{(n)} [ p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)} ] =$$

$$\varphi_{ij}^{(n)} [ p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)} ] = \varphi_{ij}^{(n)} [ \bar{p}^{(n)} ],$$

$$1 \leq i, j \leq m+1, 0 \leq \varphi_{ij}^{(n)} \leq 1, \sum_{j=1}^{m+1} \varphi_{ij}^{(n)} = 1,$$

en supposant que la  $n$ -ième expérience nous a donné la valeur  $\omega_i$ .

Les fonctions  $\varphi_{ij}^{(n)}$  peuvent être considérées comme les fonctions fondamentales de la chaîne considérée.

Il résulte de la définition que la chaîne est complètement déterminée par les matrices

$$\varphi^{(n)} = (\varphi_{ij}^{(n)}) .$$

et par la répartition initiale  $p_j$  ( $1 \leq j \leq m+1$ ).

Une chaîne à liaisons complètes est homogène si les fonctions fondamentales ne dépendent pas de  $n$ .

Notons par  $P_i^{(n)}(\bar{p})$ , ( $1 \leq i \leq m+1$ ) la probabilité que la variable  $\xi_n$  prenne la valeur  $\omega_i$ , en supposant que les probabilités des égalités  $\xi_1 = \omega_j$  soient  $p_j$ ,  $1 \leq j \leq m+1$ .

Nous allons la désigner comme probabilité totale.

La probabilité totale vérifie la relation fonctionnelle

$$P_i^{(n+1)}(\bar{p}) = \sum_{j=1}^{m+1} p_j P_i^{(n)} [ \bar{\varphi}_j(\bar{p}) ], \quad 1 \leq i \leq m. \quad (1)$$

$$[ P_{m+1}^{(n+1)}(\bar{p}) = 1 - \sum_{j=1}^m P_j^{(n+1)}(\bar{p}) ]$$

Cette équation fonctionnelle, joue le même rôle fondamental dans l'étude des chaînes à liaisons complètes que l'équation de Chapman-Smoluchovski pour les chaînes de Markoff.

La définition d'une chaîne à liaisons complètes, simple et homogène, peut être interprétée aussi de la manière suivante : la probabilité d'avoir dans la  $(n+1)$ -ème expérience les probabilités

$\varphi_{ij}(\bar{p}^{(n)})$ ,  $1 \leq j \leq m+1$ , en sachant que les probabilités respectives de  $\xi_n$  ont été  $p_i^{(n)}$ ,  $1 \leq j \leq m+1$  est égale à  $p_i^{(n)}$ .

Cette interprétation conduit à l'introduction de la chaîne associée à une chaîne à liaisons complètes, définie par les variables  $\eta_{n+1}$ , qui représentent les vecteurs de probabilités correspondant à la suite  $\xi_n$ , telle que

$$\eta_{n+1} : \begin{pmatrix} \bar{\varphi}_i(\bar{p}^{(n)}) \\ p_i^{(n)} \end{pmatrix}, \quad 1 \leq i \leq m+1$$

La suite  $\eta_{n+1}$  représente une chaîne de Markoff de nature spéciale, dont les propriétés n'ont pas été étudiées systématiquement jusqu'à présent. Les chaînes à liaisons complètes ont été étudiées pour la première fois dans les travaux [1], [2].

II. Les chaînes à liaisons complètes s'imposent comme un instrument d'étude des phénomènes de nature probabiliste dans la dépendance desquels interviennent aussi leurs probabilités de réalisation consécutive.

Supposons qu'à une certaine étape une personne possède une série des connaissances, ce qui lui permet de donner certaines réponses justes avec des probabilités connues. Jusqu'à l'étape suivante la personne est soumise à un processus pour apprendre. Les probabilités qu'elle donne de bonnes réponses dans la nouvelle étape ne sont évidemment plus les mêmes que dans l'étape précédente, mais sont certainement des fonctions de celle-ci. Les chaînes à liaisons complètes nous offrent donc tout naturellement, suivant la nature de ces fonctions, des modèles stochastiques pour apprendre.

Un exposé systématique de l'application de ces modèles et des résultats obtenus se trouve dans le livre de R. Bush et F. Mosteller [3]. On voit, d'après la nomenclature de ces auteurs, que les chaînes à liaisons complètes linéaires constituent des schémas pour des réponses contrôlés par le sujet.

Les propriétés d'ergodicité pour les chaînes à liaisons complètes ont été étudiées dans les travaux [1], [4]. Pourtant les valeurs des probabilités limites n'ont pas été déterminées. C'est à cause de cela que dans le "problème de l'apprendre" ces valeurs limite sont calculées par diverses méthodes d'approximation ou par de procédés statistiques, comme par exemple la méthode Monte-Carlo.

Dans ce travail nous nous proposons d'exposer une méthode pour trouver les valeurs des limites des probabilités  $P_i^{(n)}(\bar{p})$ , pour  $n \rightarrow \infty$ , s'ils existent.

III. Soit

$$\dots \xi_{-2}, \xi_{-1}, \xi_0, \xi_1, \dots$$

une chaîne linéaire à liaisons complètes, qui peut prendre les valeurs  $\omega_1, \omega_2, \dots, \omega_{m+1}$ .

Notons par  $p_1, p_2, \dots, p_m, p_{m+1} = 1 - \sum_{i=1}^m p_i$  les probabilités que  $\xi_n$  prenne respectivement les valeurs  $\omega_1, \omega_2, \dots, \omega_{m+1}$ , après un chemin  $c$  connu de valeurs prises par les variables antérieures à  $\xi_n$ .

Soit

$$A = (a_{ij}), \quad (i, j = 1, 2, \dots, m)$$

une matrice carrée, dont le déterminant  $|A| \neq 0$ .

Soit de même la matrice

$$\vartheta = (\vartheta_{ij}), \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, m+1).$$

Les nombres  $a_{ij}$  et  $\vartheta_{ij}$  sont réels.

Notons :

$$\bar{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix}$$

Si  $\xi_n = \omega_1$ , les probabilités  $p'_1, p'_2, \dots, p'_m$  que  $\xi_{n+1}$  prenne respectivement les valeurs  $\omega_1, \omega_2, \dots, \omega_m$  sont données par la transformation linéaire

$$T_i : \bar{p}' = A(\bar{p} - \bar{\vartheta}_i) + \bar{\vartheta}_i; \quad (i = 1, 2, \dots, m)$$

Dans cette relation on a noté

$$p' = \begin{pmatrix} p'_1 \\ p'_2 \\ \vdots \\ p'_m \end{pmatrix} \quad \vartheta = \begin{pmatrix} \vartheta_{i1} \\ \vartheta_{i2} \\ \vdots \\ \vartheta_{im} \end{pmatrix}$$

Donc  $\vartheta$  représente le vecteur donné par la  $i$ -ème colonne de  $\vartheta$ .

On suppose qu'on a toujours

$$0 < \delta < p'_k \leq \delta' \leq 1 \quad , \quad (k = 1, 2, \dots, m) \quad (I)$$

quels que soient

$$0 \leq p_k \leq 1 \quad , \quad (k = 1, 2, \dots, m).$$

Si cette condition est remplie, les probabilités totales de ces chaînes ont des valeurs limite, pour  $n \rightarrow \infty$ , qui sont indépendantes de  $A$ .

La démonstration peut résulter d'un théorème de T. E. Harris [5] ou du théorème général énoncé dans [1].

Considérons la transformation :

$$T : y' = Ty.$$

Soit  $\lambda$  une valeur propre de  $T$ ,  $z = (z_1, z_2, \dots, z_m)$  un vecteur propre correspondant à  $\lambda$  et  $z' = (z'_1, z'_2, \dots, z'_m)$  son vecteur transformé  $z' = \lambda z$ .

Notons par  $C$  l'hypercube défini par les relations :

$$0 \leq p_i \leq 1 \quad , \quad (i = 1, 2, \dots, m).$$

Si  $|\lambda| > 1$ , les itérés du point  $z \neq 0$ , sortent, après un rang quelconque de  $C$ , contrairement à l'hypothèse (I).

On a donc toujours  $|\lambda| \leq 1$ . Mais on ne peut pas avoir  $|\lambda| = 1$ . En effet si  $\lambda$  est une valeur propre sa conjuguée  $\bar{\lambda}$  est aussi une valeur propre. Dans le plan formé par les vecteurs propres qui correspondent aux  $|\lambda| = 1$ , et  $|\bar{\lambda}| = 1$ , la transformation conserve l'aire délimitée dans ce plan par  $C$ , ce qui est contraire à l'hypothèse (I).

On a donc toujours :

$$|\lambda| \leq k \leq 1 \quad (2)$$

Notons par :

$$Q_i(u) = P(\xi_n = \omega_i \mid \xi_{n-1} = u_1, \xi_{n-2} = u_2, \dots) \quad (3)$$

la probabilité que  $\xi_n = \omega_i$ , conditionnée par le fait que les variables précédentes  $\xi_{n-1}, \xi_{n-2}, \dots$  ont pris respectivement les valeurs  $u = (u_1, u_2, \dots)$  de l'ensemble  $\omega_1, \omega_2, \dots, \omega_{m+1}$ .

Considérons deux chemins

$$u = (u_1, u_2, \dots) \text{ et } v = (v_1, v_2, \dots).$$

On note

$$(u \equiv v)_s$$

si

$$u_1 = v_1, u_2 = v_2, \dots, u_s = v_s, u_{s+1} \neq v_{s+1}$$

et par

$$\varepsilon_s = \sup [Q_i(u) - Q_i(v)],$$

si

$$(u \equiv v)_s \text{ et } i = \{1, 2, \dots, m, m+1\}.$$

T. E. Harris [5] a montré que si

$$Q_i(u) \geq \Delta > 0 \quad (II)$$

$$\lim_{s \rightarrow \infty} \varepsilon_s = 0, \quad \sum_{s=0}^{\infty} \prod_{k=0}^s \left(1 - \frac{m \varepsilon_k}{2}\right) = \infty \quad (III)$$

il y a une unique chaîne à liaisons complètes stationnaire, dont les probabilités fondamentales sont données par (3).

On voit facilement que les conditions (II) et (III) sont remplies.

Si  $u = \omega_j$  et

$$R_j = P[\xi_{n-1} = \omega_j \mid \xi_{n-2} = u_2, \xi_{n-3} = u_3, \dots] \quad (j = 1, 2, \dots, m)$$

on obtient  $Q_i(u)$  par l'application de la transformation  $T_i$  aux valeurs  $R_j$ , ( $j = 1, 2, \dots, m$ ). Il en résulte de (I) que

$$Q_i(u) \geq \delta > 0$$

De la relation (1) on déduit l'inégalité

$$\epsilon_r < h k',$$

h étant une constante, qui montre que aussi la condition (III) est satisfaite.

Notons alors par

$$P_j = P(\xi_n = \omega_j), \quad (j = 1, 2, \dots, m + 1)$$

la valeur constante de la probabilité totale de la chaîne stationnaire.

On a

$$P_j = \int Q_j(u) dG(u), \quad (j = 1, 2, \dots, m),$$

G(u) étant la fonction de répartition de la variable  $u = (u_1, u_2, \dots)$  définie par les égalités

$$\xi_{n-1} = u_1, \quad \xi_{n-2} = u_2, \dots, \quad \xi_0 = u_n, \quad \xi_{-1} = u_{n+1}.$$

Considérons la fonction caractéristique de la chaîne associée

$$f(t) = \int e^{i \sum_{j=1}^m t_j q_j(u)} dG(u).$$

La chaîne étant stationnaire on a :

$$\int e^{i \sum_{k=1}^m t_k q_k(u)} dG(u) = \sum_{j=1}^{m+1} \int e^{i \sum_{k=1}^m t_k T_{jk}(q_j)} Q_j(u) dG(u).$$

Les moments du premier ordre de cette distribution existent, car ils sont respectivement égaux aux probabilités totales de la chaîne stationnaire.

En dérivant successivement par rapport aux variables  $t_i (i = 1, 2, \dots, m)$  on obtient pour  $t_i = 0, (i = 1, 2, \dots, m)$ .

$$P_k - \sum_{i=1}^{m+1} P_i \vartheta_{ik} = \sum_{j=1}^m a_{kj} (P_j - \sum_{i=1}^{m+1} P_i \vartheta_{ij}).$$

ou

$$P_k = \sum_{i=1}^{m+1} P_i \vartheta_{ik}, \quad (k = 1, 2, \dots, m).$$

Les probabilités  $P_j, (j = 1, 2, \dots, m)$  de la chaîne stationnaire ne dépendent pas de A.

On doit encore remarquer que  $\vartheta_i$  est le point fixe de la transformation  $T_i$  et que se trouve dans l'intérieur du  $C$ . Par conséquent les éléments du  $\vartheta_i$  peuvent être considérés comme probabilités.

On peut donc énoncer cette propriété : les probabilités totales de la chaîne à liaisons complètes stationnaire, définie par les transformations  $T_i$ , sont respectivement égales aux probabilités de la chaîne de Markoff, simple, stationnaire, dont les probabilités de passage sont les coordonnées des points fixes des transformations  $T_i$  ( $i = 1, 2, \dots, m$ ).

On obtient ainsi une classe de chaînes de variables aléatoires qui généralisent les chaînes de Markoff. Pour  $m = 1$ , ces chaînes remplissent la condition ( $\alpha$ ). [3].

III. Passons maintenant au cas général de chaînes à liaisons complètes linéaires.

Soit :

$$B^i = (b_{jk}^i), \quad (i, j, k = 1, 2, \dots, m+1)$$

une matrice dont les éléments  $b_{jk}^i$  sont des nombres réels, vérifiant les conditions :

$$0 < b_{jk}^i < 1 \quad (4)$$

$$\sum_{j=1}^{m+1} b_{jk}^i = 1 \quad (5)$$

Les probabilités  $\bar{p}'(p_1', p_2', \dots, p_{m+1}')$  résultent de  $\bar{p}(p_1, p_2, \dots, p_m)$  par l'application des transformations linéaires :

$$T_i : \bar{p}' = B^i \bar{p}. \quad (i = 1, 2, \dots, m+1)$$

Supposons que  $\bar{p}'$  remplisse la condition (I) et que

$$\sum_{j=1}^m |b_{jk}^i - b_{jk}^{m+1}| < 1 \quad (i, k = 1, 2, \dots, m) \quad (IV)$$

Dans ces conditions le principe ergodique est satisfait [1]. La probabilité totale  $P_i^{(n)}(\bar{p})$ , ( $i = 1, 2, \dots, m$ ) tend, pour  $n \rightarrow \infty$ , vers une limite indépendante de  $\bar{p}$ .

Pour déterminer cette limite on doit résoudre le système

$$p_i' = \sum_{j=1}^{m+1} p_j \varphi_{ji}(\bar{p}), \quad (i = 1, 2, \dots, m),$$

qui dans le cas de chaînes linéaires peut s'écrire sous une forme plus simple.

Considérons la matrice

$$B(\bar{p}) = (B^1 \bar{p} \ B^2 \bar{p} \ \dots \ B^{m+1} \bar{p})$$

où  $B^i \bar{p}$  sont aussi de matrices.

$B(\bar{p})$  a donc  $(m + i)$  lignes et  $(m + 1)$  colonnes.

Si nous notons par  $\bar{p}'$  la matrice transposée de  $\bar{p}$ , le système précédent prend la forme

$$\bar{p} = \bar{p}' B(\bar{p}). \quad (6)$$

D'après la théorie générale exposée dans [1] le système (6) a une seule solution acceptable.

#### BIBLIOGRAPHIE

- [1] ONICESCU O., MIHOC G. - Sur les chaînes de variables statistiques. Bull. Sci. Math. 59, (1935).
- [2] ONICESCU O., MIHOC G. - Sur les chaînes statistiques. C.R. Acad. Sci. Paris, 200 (1935).
- [3] BUSH R., MOSTELLER F. - Stochastic Models per Learning, New-York, (1955).
- [4] FORTET R. - Sur l'itération des substitutions algébriques linéaires a une infinité de variables et ses applications à la théorie des probabilités en chaîne. Rev. Ci. Lima. 40. (1938).
- [5] HARRIS T.E. - On chaîne of Infinite Order. Pacific J. Math. 5. (1955).

