CAHIERS DU BURO

MAURICE GIRAULT

Processus aléatoires de Markov

Cahiers du Bureau universitaire de recherche opérationnelle. Série Recherche, tome 5 (1964), p. 35-53

<http://www.numdam.org/item?id=BURO_1964__5__35_0>

PROCESSUS ALÉATOIRES DE MARKOV

par

Maurice GIRAULT

A - GÉNÉRALITÉS ET DÉFINITIONS

1/ Un système S évolue dans le temps d'une manière aléatoire. A chaque instant, il est dans l'un des états : E_1, E_2, ..., E_r et l'on s'intéresse à la suite des états pris par le système sur une suite ordonnée et discrète d'instants : $\theta_0 < \theta_1 < \theta_2 < \ldots < \theta_k \ldots$ (ces instants étant ou non aléatoires). On notera donc plus simplement ces instants par leurs rangs : 0, 1, 2, ..., k, ...

2/ Notations

L'état pris par le système à l'instant n est noté E^n (c'est, considéré à priori, un état aléatoire) et l'on pose

$$\text{Prob}\,\{E^n = E_i\} = A^n_i$$

de même

$$\text{Prob}\left\{E^{n_1} = E_{i_1}\ ;\ E^{n_2} = E_{i_2}\ \ldots\ E^{n_h} = E_{i_h}\right\} = A^{n_1\, n_2\, \ldots\, n_h}_{i_1\, i_2\, \ldots\, i_h}$$

pour $$n_1 < n_2 < \ldots < n_h$$

Le processus est défini en probabilité par l'ensemble des fonctions A pour toutes les valeurs possibles de (n) et de (i) en nombre (h) quelconque. Ces données sont évidemment surabondantes et ne sont pas indépendantes les unes des autres. Si l'ensemble des instants où l'on observe le processus était fini ($\{\theta_1\ \theta_2 \ldots\ \theta_k\}$ par exemple), chaque réalisation du processus se traduirait par un élément aléatoire à k dimensions et il suffirait de se donner les r^k probabilités suivantes : (qui ne sont liées que par la seule relation : $\Sigma\,(A) = 1$)

$$\text{Prob}\left\{E^1 = E_{i_1} \ ; \ E^2 = E_{i_2} \ ; \ \dots \ ; \ E^k = E_{i_k}\right\} = A_{i_1 i_2 \dots i_k}$$

$$\text{où} \quad i_1 = 1, 2, \dots, r$$

$$i_2 = \dots\dots\dots\dots$$

$$i_k = \dots\dots\dots$$

car on peut évidemment exprimer toutes les lois conditionnelles à partir des (A) précédents.

On ne peut pas conserver cette représentation lorsque k tend vers l'infini ; car, en général, chacune des probabilités précédentes tend vers 0. (les évènements élémentaires de cette classification sont tous de probabilité nulle).

3/ Représentation par développement temporel

Il est alors commode de se donner :

a) *la loi initiale* : par l'ensemble des r probabilités

$$\{A_1^\circ, A_2^\circ, \dots, A_r^\circ\}$$

qu'on peut considérer comme composantes du vecteur $\vec{P}_\circ$.

b) *la loi d'évolution du système*

La loi d'évolution entre les instants 0 et 1 est définie par la donnée des r^2 probabilités suivantes

$$p_{i_1}^{(i_\circ)} = \text{Prob}\left\{E^1 = E_{i_1}/E^\circ = E_{i_\circ}\right\} = \frac{A_{i_\circ\, i_1}^{\circ,\,1}}{A_{i_\circ}^\circ} \quad \text{pour} \quad \begin{cases} i_\circ = 1, 2 \dots r \\ i_1 = 1, 2 \dots r \end{cases}$$

Appelons $P_1^{(\circ)}$ le tableau de ces nombres.

De même $P_2^{(\circ,1)}$ désignera l'ensemble des r^3 probabilités des états E^2 connaissant $E^\circ = E_{i_\circ}$ et $E^1 = E_{i_1}$

$$p_{i_2}^{(i_\circ i_1)} = \frac{A_{i_\circ i_1 i_2}^{\circ,1,2}}{A_{i_\circ\, i_1}^{\circ\ \ 1}} \quad \text{pour} \quad \begin{cases} i_\circ = 1, 2 \dots r \\ i_1 = \text{"} \quad \text{"} \\ i_2 = \text{"} \quad \text{"} \end{cases}$$

Plus généralement $P_{k+1}^{(o,\ 1,\ \dots\ k)}$ désigne l'ensemble des r^{k+1} probabilités conditionnelles des états à l'instant $(k + 1)$ connaissant les états pris aux instants 0, 1, ... k.

Le système

$$\mathrm{I} \quad \left\{ \begin{array}{l} P_1^{o} \\ P_2^{(o1)} \\ P_3^{(o12)} \end{array} \right.$$

est dit *"système des lois conditionnelles fondamentales"*.

4/ Suite de Markov

Les processus ainsi définis, malgré leurs caractères discrets (suite discrète d'instants) et *finis* (ensemble fini des états possibles) sont encore trop généraux pour qu'on puisse en dire plus sans spécifier leurs lois.

On obtient un type très intéressant de processus en posant la condition suivante :

4.1 - CONDITION DE MARKOV :

Pour toute suite ordonnée d'instants :

$$n_1 < n_2 < \dots < n_k < n_{k+1} < \dots < n_{k+s}$$

la loi conditionnelle de l'ensemble des états $(E^{n_{k+1}}, E^{n_{k+2}}, \dots, E^{n_{k+s}})$ connaissant $\left\{ E^{n_1} = E_{i_1}\ ;\ E^{n_2} = E_{i_2}\ ;\ \dots\ ;\ E^{n_k} = E_{i_k} \right\}$ ne dépend que du *dernier instant* n_k et et du dernier état E_{i_k} pris par le système sur l'ensemble donné. En d'autres termes, si l'on connaît l'histoire du système jusqu'à l'instant actuel : son état dans le présent résume toute l'information utile pour connaître son comportement futur.

4.2 - Le système I se simplifie beaucoup et la représentation temporelle devient la suivante :

a) *loi initiale*, vecteur $\vec{P}_o$ de composantes $A_1^o\ A_2^o\ \dots\ A_r^o$

b) *lois conditionnelles des états*

b$_1$) *Lois à l'instant (1) connaissant l'état* E^o

Nous noterons maintenant : $p_{ji}^{o,1} = \text{Prob}\left\{E^1 = E_i /_{E^o} = E_j\right\}$

et $P_1^o = \pi_1^{(o)}$ = matrice $\left\| p_{j,i}^{o,1} \right\|_{r \times r}$ où $\begin{cases} i = \text{indice de ligne} \\ j = \text{" de colonne} \end{cases}$

chaque colonne de cette matrice est un vecteur de probabilités (la somme de ses composantes = 1)

b$_2$) *Lois à l'instant (2) connaissant l'état* E^1

Puis $P_2^{(o,1)} = \pi_2^{(1)}$ = matrice $\left\| p_{j,i}^{1,2} \right\|$

etc ... $P_{k+1}^{(o,1 \ldots k)} = \pi_{(k+1)}^{(k)}$ = matrice $\left\| p_{j,i}^{(k,\ k+1)} \right\|$

Le système des lois conditionnelles fondamentales devient :

$$\text{II} \left\{ \pi_1^{(o)} \; ; \; \pi_2^{(1)} \ldots \pi_{k+1}^{(k)} \right.$$

Ainsi l'évolution du système est complètement déterminée par la donnée des probabilités conditionnelles des états pris en deux instants θ consécutifs. Pour exprimer cette idée, on a appelé ce processus "chaîne de Markov" (ou processus en chaîne de 1 *). Actuellement le mot "chaine" a en algèbre un sens précis, totalement étranger à l'idée précédente ; et, comme l'étude des processus de Markov fait intervenir l'Algèbre, il semble préférable de désigner par "suites de Markov" les suites temporelles discrètes aléatoires obéissant à la condition (4.1).

5/ Le système (II) definit en probabilité le processus

5.1 - LOIS A PRIORI

Désignons par $\vec{P}_n$ le vecteur matrice colonne de composantes

----

* Dans un processus en chaîne de 2, la loi conditionnelle des états futurs serait fonction des deux derniers états connus et des deux derniers instants correspondants.

$$\begin{bmatrix} p_n^1 \\ p_n^2 \\ \cdot \\ \cdot \\ \cdot \\ p_n^r \end{bmatrix}$$

où $p_n^{(i)} = \text{Prob}\,\{E^n = E_{(i)}\}$ Ce vecteur exprime la loi de probabilité du système à l'instant (n). Or (en notations matricielles) on a :

$$\boxed{\vec{P}_n = \pi_n^{(n-1)}\ \pi_{n-1}^{(n-2)} \cdots \pi_2^{(1)}\ \pi_1^{(o)}\ \vec{P}_o} \qquad (5\text{-}1)$$

5.2 - LOIS CONDITIONNELLES

La loi conditionnelle des états à l'instant $(n + k)$ connaissant l'état $E_{(i)}$ à l'instant (n), et cela en fonction de $E_{(i)}$, est définie par la matrice $(r \times r)$ de terme général :

$$m_{i,j} = p_{j,i}^{n,n+k} = \text{Prob}\left\{E^{(n+k)} = E_i/E^n = E_j\right\} \qquad \begin{array}{l} i = 1,\ 2,\ \ldots,\ r \\ j = 1,\ 2,\ \ldots,\ r \end{array}$$

Soit $\pi_{n+k}^{(n)}$ cette matrice.

Toutes ces matrices sont liées par des relations telles que :

$$\pi_k^{(h)} = \pi_k^{(s)}\ \pi_s^{(h)} \qquad \text{si} \quad h < s < k \qquad (5\text{-}2)$$

Cette relation (dite de Chapman Kolmogorov) n'exprime pas autre chose que la définition des probabilités conditionnelles.

En particulier

$$\pi_{n+k}^{(n)} = \pi_{n+k}^{(n+k-1)}\ \pi_{n+k-1}^{(n+k-2)} \cdots \pi_{n+1}^{(n)}$$

6/ Suite de Markov homogène

Pour compléter l'étude, sans spécifier les lois il faut faire de nouvelles hypothèses. La plus simple consiste à poser :

$$\pi^{(k)}_{k+1} = M \quad \text{matrice indépendante de } k.$$

Par suite $\pi^{(n)}_{n+k} = (M)^k$ puissance d'ordre k de M.

L'évolution du système entre deux instants (n) et $(n + k)$ ne dépend que de k ; elle reste inchangée par une translation sur l'échelle des temps ; c'est ce que l'on exprime en disant que le processus est *homogène* (dans le temps).

Le système est alors défini en probabilité par les seules données de $\vec{P}_0$ (loi initiale) et de M (matrice élémentaire des probabilités conditionnelles dite aussi "matrice des probabilités de passage").

La loi de probabilité à l'instant (n) est $\vec{P}_n = M^n \vec{P}_0$

et $\pi^{(n)}_{n+k} = M^k$

L'étude de l'évolution du système se ramène à celle des puissances de M et le comportement à long terme est défini par $\lim\limits_{k=\infty} [M^k]$. Cette évolution est particulièrement simple lorsque la matrice M ne présente pas certaines particularités que nous préciserons plus loin. On est toujours dans ce cas (dit "régulier") si tous les termes de la matrice M sont positifs.

B - ETUDE BASÉE SUR LA CONVEXITÉ

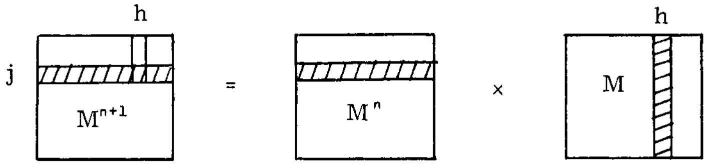
7/ Limite pour $k = \infty$ de M^k lorsque M a tous ses termes positifs

Soit $M = \text{matr}\,(p_{j,i})_{r \times r}$

où $$p_{j,i} = \text{Prob}\left\{E^{(k+1)} = E_i / E^{(k)} = E_j\right\}$$

ou probabilité de passer de l'état E_j à l'instant (k) à l'état E_i à l'instant suivant $(k + 1)$.

et $M^n = \text{matr}\,[P^{(n)}_{j,i}]_{r \times r}$ où $P^{(n)}_{j,i} = \text{Prob}\ \ E^{k+n} = E_i / (E^k = E_j)$



Un terme $P_{hj}^{(n+1)}$ de la ligne (j) de M^{n+1} s'obtient comme suit :

$$P_{h,j}^{(n+1)} = p_{n1}\, P_{1j}^{(n)} + p_{h2}\, P_{2j}^{(n)} + \dots + p_{hr}\, P_{rj}^{(n)} \tag{1}$$

Or $(p_{h1}\ ;\ p_{h2} \dots p_{hr})$ sont des coefficients non négatifs de somme 1. Tout terme de la ligne j de M^{n+1} est barycentre des termes de la même ligne de M^n. L'amplitude λ_n^j des valeurs des termes de la ligne j de M^n diminue régulièrement avec (n). Cette amplitude tend vers zéro, comme nous allons le voir, si tous les termes de M sont positifs. Dans ces conditions, pour n très grand les termes d'une même ligne sont sensiblement égaux.

Supposons que

$$\forall i, \qquad \forall j \qquad : \ p_{ij} \geq \varepsilon > 0$$

Ecrivons les $\{P_{ij}^{(n)}\}_{i = 1,\, 2,\, \dots\, r}$ par ordre croissant :

$$0 \leq y_1 \leq y_2 \dots \leq y_r \leq 1$$

$$y_1 = \underline{p}_j^{(n)} \text{ est le minimum des } \{P_{1j}^{(n)}\}_i$$

et

$$y_r = \overline{p}_j^{(n)} \text{ est le maximum du même ensemble.}$$

et posons

$$\lambda_n^j = y_r - y_1$$

La relation (1) s'écrit :

$$z_h = y_1\, c_h^1 + y_2\, c_h^2 + \dots + y_r\, c_h^r \tag{2}$$

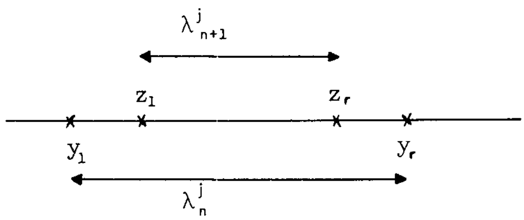
on a posé $P_{hj}^{n+1} = z_h$ les $\{c_h^i\}_i$ sont les coefficients $\{P_{hi}\}_i$ dans un certain ordre.

$$z_h - y_1 = (y_1 - y_1)\, c_h^1 + (y_2 - y_1)\, c_h^2 + \dots + (y_r - y_1)\, c_h^r$$

termes tous ≥ 0 le dernier est au moins égal à $\lambda_n^j\, \varepsilon$ d'où

$$z_h - y_1 \geq \varepsilon\, \lambda_n^j$$

De même en formant $y_r - z_h$ on démontre que $y_r - z_h \geq \varepsilon \lambda_n^j$.



On en déduit facilement la relation $\lambda^j_{n+1} \leq \lambda^j_n (1-2\varepsilon)$

La suite des amplitudes $\{\lambda^j_n\}_n$ décroit plus vite qu'une progression géométrique de raison $1-2\varepsilon < 1$, la limite pour $n = \infty$ est donc nulle et par suite la matrice M_n tend vers une matrice M^∞ ayant toutes ses colonnes identiques

$$P^\infty_{hj} = P^\infty_{kj} = a_j \quad \forall h, k, j$$

Ces probabilités limites a_j sont toutes différentes de zéro puisque $p^{(n+1)}_{hj}$ est compris entre $\overline{p}^{(n)}_j$ et $\underline{p}^{(n)}_j$ eux-mêmes compris entre $\overline{p}^{n-1}_j$ et $\underline{p}^{n-1}_j$ etc. (en remontant, de proche en proche) compris entre $\overline{P}^1_1$ et $\underline{P}^1_1$ qui sont des valeurs $p_{i1} \ldots$ et p_{j1} de M supposées $\neq 0$.

D'où la conclusion

Si M = matrice $(p_{ji})_{r \times r}$ a tous ses termes $\geqslant \varepsilon > 0$

la matrice M^n tend vers une matrice M^∞ de la forme

$$M^\infty = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} A \end{bmatrix} \cdots \begin{bmatrix} A \end{bmatrix} \quad \text{où} \quad A = \begin{bmatrix} a_1 \\ a_2 \\ a_r \end{bmatrix}$$

est un vecteur de probabilités à composantes positives
$a_1 > 0$
$a_1 + a_2 + \ldots + a_r = 1$

8/ Résultat fondamental

On dit qu'on est dans le cas *"positivement régulier"* si, lorsque n tend vers l'infini, la matrice $(M)^n$ tend vers une matrice composée de r colonnes A identiques : dont les composantes $(a_1\ a_2 \dots a_r)$ sont toutes positives.

Dans ces conditions quelles que soient les conditions initiales (loi $\vec{P}_0$) la loi limite pour $n = \infty$ de $E^{(n)}$ est $\vec{A}$

En effet $$\vec{P}_n = M^n P_0$$

les limites pour $n = \infty$ sont :

$$P_\infty = \begin{bmatrix} \begin{bmatrix} A \end{bmatrix} & \begin{bmatrix} A \end{bmatrix} & \dots & \begin{bmatrix} A \end{bmatrix} \end{bmatrix} \times \begin{bmatrix} P_0 \end{bmatrix} = A$$

Remarques 1/ L'évolution de M_n vers sa limite a le caractère exponentiel ; elle est donc en général rapide ;

2/ C'est la loi *à priori* (ou marginale) de E^n qui, pour n assez grand, est $\vec{A}$, mais les lois conditionnelles (de E^{n+r}/E^n par exemple) restent décrites par la matrice M et ses puissances successives.

9/ Restriction des hypothèses

On a supposé jusqu'ici que tous les termes p_{ji} de M étaient différents de zéro. On obtient les mêmes résultats s'il existe un entier s tel que M^s ait tous ses termes non nuls.

Cette dernière propriété sera réalisée pour M^{s+1} donc pour $M^n\ \forall\ n > s$.

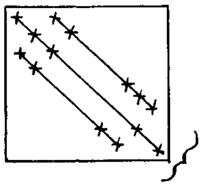
D'autre part la suite des matrices $M^s\ (M^s)^2 \dots (M^s)^n$ converge vers :

$$\begin{bmatrix} A \end{bmatrix} \dots \begin{bmatrix} A \end{bmatrix} = \mathfrak{A}$$

or $$\vec{P}_{ns+h} = M^{ns}\, M^h\, \vec{P}$$

donc quelque soit h, si n tend vers l'infini $\vec{P}_{ns+h}$ tend vers $\vec{A}$.

Les conditions précédentes sont réalisées si la matrice M n'a aucun élément nul dans sa diagonale principale, ainsi que dans les deux rangées obliques voisines (v. fig.). On constate en effet qu'en prenant les puissances successives de M le nombre de rangées à termes tous positifs augmente de 2 unités chaque fois.



C - ETUDE PAR ANALYSE SPECTRALE

10/ Valeurs propres et vecteurs propres

Si $\vec{P}_n = M \vec{P}_{n-1} = M^n \vec{P}_0$ tend vers une limite pour $n = \infty$ cette limite $\vec{X}$ réalise

$$\vec{X} = M \vec{X} \qquad \text{avec} \qquad \vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_r \end{bmatrix}$$

L'équation matricielle (1) résume un système de r équations linéaires à r inconnues. Pour le discuter il est commode de généraliser le problème et de chercher les vecteurs $\vec{X}$ tels que $\vec{X}$ et $M\vec{X}$ aient même support, c'est-à-dire tels que :

$$M\vec{X} = x\vec{X}$$

Si I est la matrice unitaire $r \times r$:

$$M\vec{X} = s\,I\,\vec{X} \quad \text{soit} \quad (M - s\,I)\,\vec{X} = 0.$$

Pour que le système (2) ait des solutions non nulles, il faut et il suffit que

$$\det\,(M - s\,I) = 0$$

les racines s de l'équation $\det(M - sI) = 0$ (10) sont dites *valeurs propres* (ou valeurs spectrales) de M. L'ensemble des valeurs propres s'appelle le spectre de la matrice M. Si s_i est une valeur propre : $(M - sI)\vec{X} = 0$ a une solution au moins non nulle : soit $\vec{V}_i$, dit *vecteur propre* associé à la valeur propre s_i.

11/ Particularités des valeurs propres d'une matrice de probabilités

La matrice M des probabilités conditionnelles est composée de colonnes : $A_1\ A_2 \dots A_r$ qui sont des vecteurs de probabilités

$$\vec{A}_j = \begin{bmatrix} p_{j1} \\ p_{j2} \\ p_{jr} \end{bmatrix} \quad \text{où} \quad 0 \leq p_{ji} \leq 1 \quad \text{et} \quad p_{j1} + p_{j2} + \dots + p_{jr} = 1$$

C'est là une particularité importante des matrices M qui entraîne certaines conséquences sur $\{s_i\}$.

Tout d'abord : si s appartient au spectre de M : le déterminant

$$\det(M - sI) = 0$$

mais il en est de même de : $\det(M' - sI) = \det(M - sI)'$ et réciproquement* et le système :

$$s\begin{bmatrix} \vec{Y} \end{bmatrix} = \begin{bmatrix} \boxed{\quad A_1 \quad} \\ \boxed{\quad A_2 \quad} \\ \text{-----------} \\ \boxed{\quad A_r \quad} \end{bmatrix} \begin{bmatrix} Y \end{bmatrix} \quad \text{est également compatible}$$

----

* On désigne par M' la matrice transposée de M, c'est-à-dire obtenue à partir de M par permutation lignes-colonnes ($m'_{ji} = m_{ij}$).

Explicité il s'écrit :

$$\begin{cases} s\,Y_1 = p_{11}\,Y_1 + p_{12}\,Y_2 + \dots + p_{1r}\,Y_r \\ s\,Y_2 = p_{21}\,Y_1 + p_{22}\,Y_2 + \dots + p_{2r}\,Y_r \\ \dots\dots\dots\dots\dots\dots \\ s\,Y_r = p_{r1}\,Y_1 + p_{r2}\,Y_r + \dots + p_{rr}\,Y_r \end{cases}$$

11.1 - LES VALEURS SPECTRALES (s) ONT UN MODULE INFERIEUR OU EGAL A L'UNITE

Soit R le module maximum des composantes $Y_1\ Y_2 \dots Y_r$ d'un vecteur propre associé à (s), maximum atteint par exemple par Y_j. On a :

$$s\,Y_j = p_{j1}\,Y_1 + p_{j2}\,Y_2 + \dots + p_{jr}\,Y_r$$

où les p_{ji} sont ≥ 0 et de somme unitaire $s\,Y_j$ est le barycentre des $\{Y_i\}$ affectés des coefficients $\{p_{ji}\}$

$$|s|\,|Y_j| \leq \sum_k p_{jk}\,|Y_k| \leq R$$

$$|s|\,R \leq R \qquad \text{donc} \qquad \underline{|s| \leq 1}$$

11.2 - PARMI LES VALEURS PROPRES, FIGURE TOUJOURS LA VALEUR $s = 1$

Pour $s = 1$ l'équation (10) s'écrit $\det(M - 1) = 0$ ce qui est manifestement réalisé car la somme des termes de chaque colonne est nulle.

11.3 - LES VALEURS PROPRES COMPLEXES DE MODULE UNITAIRE SONT DES RACINES ENTIERES DE L'UNITE

Soit s_i telle que $|s_i| = i$ et $\vec{Y} = [Y_1, Y_2, \dots Y_r]$ un vecteur propre associé à s_i par la matrice M' :

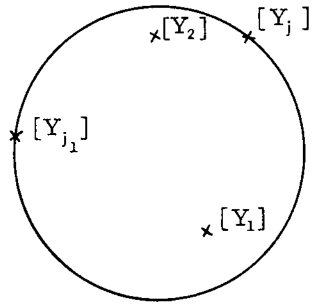
$$s_i\,\vec{Y} = M'\,\vec{Y}' \qquad (1)$$

et soit enfin R le module maximum des composantes de $\vec{Y}$:

$$R = |Y_j| \text{ par exemple}$$

La relation (1) donne pour le jème composante

$$s_i \, Y_j = \sum_k p_{jk} \, Y_k$$



Soit $[Y_k]$ le point du plan complexe d'affixe Y_k, $[s\,Y_j]$ est un point du cercle C (de centre 0, de rayon R) $[s\,Y_j]$ est barycentre des points $[Y_1] \ldots [Y_j] \ldots [Y_r]$ tous dans C ou sur sa frontière.

Donc $[s\,Y_j]$ (différent de $[Y_j]$ puisque $s \neq 1$) est l'un des points $[Y_i]$ soit $[Y_{j_i}]$ et dans l'opération convexe qui définit Y_{j_i} :

$$Y_{j_i} = s\,Y_j = \sum_k p_{jk} \, Y_k$$

les p_{jk} associés aux points autres que $[Y_{j_1}]$ sont tous nuls ; la somme des coefficients affectés à $[Y_{j_1}]$ est 1 (plusieurs valeurs Y_i peuvent être égales à Y_{j_1})

d'où finalement : il existe une composante Y_{j_1} de $\vec{Y}$ différente de Y_j et telle que

$s_i \, Y_{j_1} = Y_{j_2}$ avec $|Y_j| = |Y_{j_1}| = R$ De même

$s_i \, Y_{j_2} = Y_{j_3}$

Or il y a au plus r valeurs Y_i distinctes. Il existe donc une puissance entière h de s_i telle que

$$(s_i)^h \, X_j = X_j \quad \text{soit} \quad (s_i)^h = 1 \quad \text{avec} \quad h \leq r$$

12/ Forme canonique

12.1 - Supposons tout d'abord que les valeurs propres $\{s_i\}$ soient toutes des racines simples de l'équation det $(M - s\,I) = 0$.

Soient $s_1 = 1$; $s_2 \ldots s_r$ ces racines ordonnées par exemple suivant les modules décroissants.

On montre que la matrice M peut s'écrire

$$M = S \begin{vmatrix} s_1 & & & & 0 \\ & s_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & s_r \end{vmatrix} S^{-1}$$

où S est une matrice $r \times r$ dont les colonnes sont r vecteurs propres associés respectivement à $s_1, s_2, \ldots, s_r$.

S^{-1} est la matrice inverse de S.

La matrice intermédiaire est une matrice diagonale dite réduite de Jordan de M. On voit immédiatement que M^n vaut :

$$M^n = S \begin{vmatrix} s_1^n & & & & \\ & s_2^n & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & s_r^n \end{vmatrix} S^{-1}$$

où $s_1^n = 1$

pour $i \neq 1$ s_i^n tend vers 0 avec $1/n$

$$M^\infty = \begin{bmatrix} V_1 \end{bmatrix} \begin{bmatrix} V_1 \end{bmatrix} \ldots \begin{bmatrix} V_1 \end{bmatrix}$$

où V_1 = première colonne de S est le vecteur propre associé à $s_1 = 1$

12.2 - CAS GENERAL

Si les valeurs propres sont a, b, c, etc. à des ordres quelconques, on peut encore écrire M sous forme réduite ; mais alors la matrice intermédiaire est par exemple de la forme :

$$\begin{pmatrix} a & 1 & & & & & & & & \\ & a & 1 & & & & & & & \\ & & a & & & & & & & \\ & & & a & & & & & & \\ & & & & b & 1 & & & & \\ & & & & & b & & & & \\ & & & & & & b & & & \\ & & & & & & & c & 1 & \\ & & & & & & & & c & \\ & & & & & & & & & d \end{pmatrix}$$

si a est d'ordre 4
si b " " 3
c " " 2
les termes non écrits sont tous nuls.

La forme réduite n'est pas complètement définie par la donnée des racines, même munies de leur ordre de multiplicité ; il faut avoir la "caractéristique de Segre" ici égale à (3,1) (2,1), (2), (1) ou toute autre information équivalente.

On voit ainsi que le cas général est assez complexe, il est plus simple de le décrire et de l'étudier par des considérations algébriques sur l'ensemble des états ; ce qui constitue la *"méthode directes"*.

D - ETUDE ALGEBRIQUE DIRECTE

Nous nous plaçons toujours dans l'hypothèse d'une suite de Markov homogène dans le temps.

13/ Classification des états

Soit $E = \{E_1 \; ; E_2 \; ; \ldots \; ; E_r\}$ l'ensemble des états possibles.

13.1 - RELATION DE PREORDRE SUR LES ETATS

Un état E_j est dit *conséquent de* E_i s'il existe un entier naturel n au moins tel que

$P_{ij}^{n} > 0$ (Probabilité non nulle de passer de E_i à E_j en n étapes).

On notera cette relation :

$$E_i \longrightarrow E_j$$

Cette relation est *transtitive :*

$$\text{Si} \quad \left\{ \begin{array}{l} E_i \longrightarrow E_j \\ E_j \longrightarrow E_k \end{array} \right. \quad \Longrightarrow \quad E_i \longrightarrow E_k$$

En effet : il existe n_1 tel que $P_{ij}^{n_1} > 0$ et n_2 " " $P_{jk}^{n_2} > 0$ } la probabilité de passer de E_i à E_k en $(n_1 + n_2)$ étapes est au moins égale à $P_{ij}^{n_1}\ P_{jk}^{n_2}$: elle est donc positive.

La relation binaire ainsi définie n'est pas réflexive : pour simplifier l'exposé, on peut étendre cette relation :

On dira que E_j est *strictement conséquent* de E_i (notation $E_i \longrightarrow E_j$) s'il existe un entier n *positif* tel que $P_{ij}^n > 0$ et l'on dira que E_j est *conséquent* (sens large) de E_i (notation $E_i \longrightarrow E_j$) s'il existe un entier n *positif ou nul* tel que $P_{ij}^n > 0$.

Pour $n = 0$ en conviendra de poser $P_{ii}^0 = 1$ et $P_{ij}^0 = 0$ pour $j \neq i$. La relation $\longrightarrow$ est une relation de préordre sur les états.

13.2 - RELATION D'EQUIVALENCE CLASSES

La relation $E_i \rightleftarrows E$ qui exprime que E_i et E_j sont mutuellement conséquents l'un de l'autre est *symétrique et transitive :* c'est une relation d'équivalence qui permet de réduire en classes les éléments de E.

En effet : Soit E_i un état de E :

Rassemblons les états de E en relation d'équivalence avec E_i : par exemple E_{i_1}, E_{i_2} ... . Ce sous-ensemble de E constitue une *classe* : $\mathcal{C}_i = \{E_i, E_{i_1}, E_{i_2} \ldots\}$. Les états de $\mathcal{C}_i$ jouissent des deux propriétés suivantes :

- Deux états quelconques de $\mathcal{C}_i$ sont en relation d'équivalence entre eux : $E_{i'} \rightleftarrows E_{i''}$

- Il n'existe pas d'état n'appartenant pas à $\mathcal{C}_i$ qui soit en relation d'équivalence avec un état de $\mathcal{C}_i$.

Naturellement, la classe de E_i peut ne contenir que E_i lui-même.

Ainsi, chaque état appartient à 1 classe et à 1 seule. Le nombre de classes est au plus égal à r.

Cette répartition en classes, ainsi que toutes les considérations qui vont suivre sont basées sur la notion d'état conséquent d'un autre et cette relation binaire entre états ne dépend que de la place des 0 et des termes $\neq 0$ de la matrice M et absolument pas des valeurs numériques des termes non nuls et c'est pourquoi cette méthode algébrique est intéressante : en effet elle néglige délibérément les éléments qui n'interviennent pas dans la description de l'évolution du système.

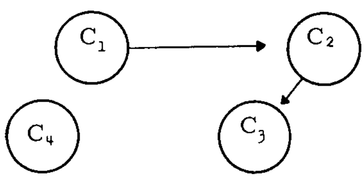
14/ Relation binaire sur les classes : classes de passage et classes finales

Soit deux classes $\mathcal{C}_i$ et $\mathcal{C}_j$ telles qu'il existe un état de $\mathcal{C}_j$ conséquent d'un état $\mathcal{C}_i$.

On en déduit immédiatement que tout état de $\mathcal{C}_j$ est conséquent de tout état de $\mathcal{C}_i$ et l'on dira que $\mathcal{C}_j$ est conséquent de $\mathcal{C}_i$; on notera $\mathcal{C}_i \longrightarrow \mathcal{C}_j$.

Si $\mathcal{C}_i \longrightarrow \mathcal{C}_j$ on ne peut avoir $\mathcal{C}_j \longrightarrow \mathcal{C}_i$ car alors tous les états de $\mathcal{C}_i$ et de $\mathcal{C}_j$ pris deux à deux seraient mutuellement conséquents : ils formeraient donc une seule et même classe.

Plus généralement, entre les classes, la relation $\longrightarrow$ établit un réseau ne contenant *aucun cycle :*



14.1 - Une classe $\mathcal{C}_i$ est dite *classe de passage* s'il existe au moins une classe conséquente de $\mathcal{C}_i$ [donc si un état au moins de $\mathcal{C}_i$ admet au moins un conséquent n'appartenant pas à $\mathcal{C}_i$].

En d'autres termes : une classe est de passage s'il est possible d'en sortir ; et, d'après ce que nous venons de voir, il sera alors impossible d'y entrer de nouveau.

14.2 - Au contraire une classe est dite *finale* si elle n'admet aucune classe conséquente : lorsqu'on a pu y entrer, il est impossible d'en sortir. Il existe toujours au moins une classe finale (conséquence du fait qu'il ne peut pas exister de cycle sur le réseau entre classes) et la probabilité d'aboutir à une classe finale est égale à 1.

14.3 - UNE CLASSE FINALE EST *CYCLIQUE* OU *ACYCLIQUE*

Elle est dite cyclique lorsque les états qui la composent peuvent être répartis en sous classes $S_1\ S_2 \dots S_h$ telles que

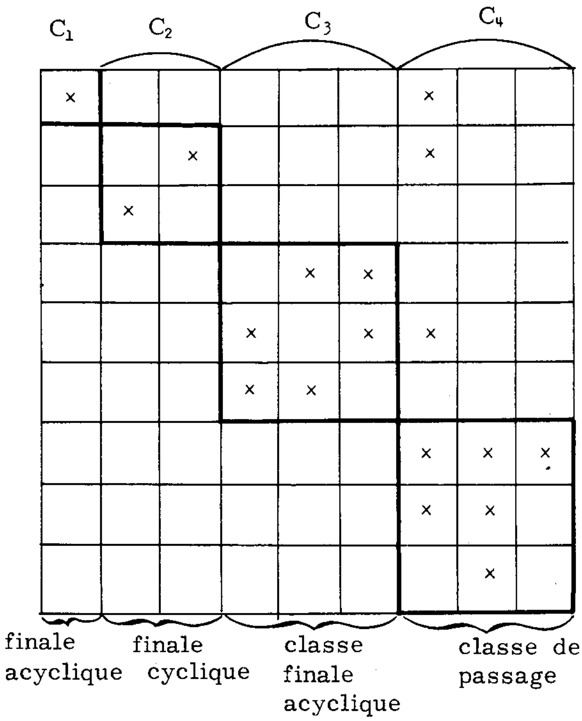
$$\text{si} \quad E^n \in S_i \Longrightarrow E^{n+1} \in S_{i+1} \qquad \text{avec la probabilité 1}$$

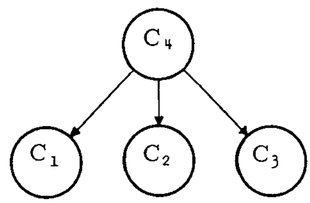
avec évidemment $S_{h+1} = S_1$

Si une classe ne comprend qu'une seule sous-classe S, elle est dite acyclique.

14.4 - EXEMPLE

Décomposition en classes de l'ensemble des états :





réseau de connexion des classes

Conclusion : Récapitulation des différentes évolutions possibles

a) Le cas le plus simple est le cas "positivement régulier" dans lequel il n'y a qu'une *seule classe* (nécessairement finale) et *acyclique*. Pour n assez grand, la matrice M^n n'a aucun terme nul. Pour n très grand, les états possibles de E^n sont tous les états considérés à priori $\{E_1, E_2, \dots, E_r\}$ et la loi de probabilité sur ces états est indépendante à la fois de n et des conditions initiales (aléatoires ou déterminées).

b) Lorsqu'il n'y a qu'une seule classe finale (avec ou sans classe de passage) et que cette classe est *acyclique*, on dit qu'on est dans le cas régulier : il n'y a qu'une seule valeur propre de module 1 : c'est $s = 1$. Pour n très grand, les états possibles de E^n sont les états de la classe finale et la loi sur ces états est indépendante à la fois de n et des conditions initiales.

c) Une classe finale *cyclique* : le spectre comprend alors c valeurs propres de module 1 dont une seule est 1 ; elles sont les racines de $Z^c - 1 = 0$. Pour n très grand, la loi de E^n peut dépendre des conditions initiales, de plus elle dépend toujours de n. Par contre, si l'on pose : $n = mc + n_1$ avec $n_1 < c$ la loi de E^n dépend de n_1 mais non de m.

d) S'il y a f classes finales, la valeur propre $s = 1$ est racine d'ordre f de l'équation $\det(M - sI) = 0$. Lorsque n augmente le processus atteint, au bout d'un nombre *fini* d'étapes l'une des classes finales et il y reste. La loi de probabilité de la classe (aléatoire) atteinte dépend de M et des conditions initiales. Par contre, à l'intérieur de la classe finale atteinte (acyclique ou cyclique) le processus évolue comme en (a) ou en (c).