Enfin nous admettrons le résultat suivant : si une suite converge par chacun des deux procédés de Cesaro et d'Euler alors les deux limites sont les mêmes.

6.3.- THEOREME : Soit P la matrice des transitions d'une chaîne ergodique (périodique ou régulière) alors :

1) la suite des matrices P^n converge au sens d'Euler vers une matrice A dont toutes les lignes sont égales à un même vecteur ligne $\vec{\alpha} = (a_1,\dots,a_N)$ et $\sum_{i=1}^{N} a_i = 1$.

2) Quel que soit la loi initiale $\vec{c}$, $\vec{c}.P^n$ converge au sens d'Euler vers $\vec{\alpha}$ quand $n \to +\infty$.

3) $\vec{\alpha}$ est l'unique solution de l'équation $\vec{\alpha}P = \vec{\alpha}$

4) $PA = AP = A$

5) la suite des matrices P^n converge aussi au sens de Cesaro vers A.

Démonstration : (*)

1) Considérons la matrice $\tilde{P} = kI+(1-k)P$ (k fixé $0 < k < 1$). Alors $\tilde{P}$ est une matrice des transitions d'une certaine chaîne de Markov. Les coefficients de $\tilde{P}$ sont > 0 aux endroits ou P avait des coefficients > 0 ; il en résulte que $\tilde{P}$ est la matrice d'une chaîne ergodique. De plus la diagonale de $\tilde{P}$ est composée d'éléments > 0 ; on peut donc retourner en i au pas suivant donc $d = 1$ et $\tilde{P}$ est une matrice régulière. D'après 5.2. on sait que $(\tilde{P})^n$ converge vers une matrice A dont toutes les lignes sont égales à $\vec{\alpha} = (a_1,\dots,a_N)$ (avec $\sum_{i=1}^{N} a_i = 1$). Ainsi $A = \lim_{n\to\infty} (kI+(1-k)P)^n$

$$\text{ie :} \qquad A = \lim_{n\to\infty} \sum_{i=0}^{n} C_n^i\, k^{n-i}(1-k)^i P^i \quad (*)$$

A est donc la limite au sens d'Euler de la suite P^n.

2) Si on multiplie à gauche l'égalité (*) par le vecteur ligne $\vec{c}$ on obtient de suite le résultat.

3) On sait déjà d'après le théorème 5.2. que $\vec{\alpha}$ est l'unique solution de $\vec{\alpha}\tilde{P} = \vec{\alpha}$. Or il est immédiat de voir que $\vec{c}P = \vec{c}$ équivaut à $\vec{c}\tilde{P} = \vec{c}$, d'où le résultat de 3).

(*) La démonstration n'est pas exigée. Le résultat peut être admis

4) même démonstration qu'en 5.2.

5) Si $n = kd$ alors partant de i, après n pas on est dans la classe cyclique de i et si k est suffisamment grand on peut être dans n'importe quel état de la classe cyclique de i. Ainsi on peut considérer P^d comme la matrice des transitions d'une chaîne de Markov ayant d ensembles ergodiques distincts et réguliers. Autrement dit on a :

$$P^d = \begin{pmatrix} \boxed{P_1} & & 0 \\ & \boxed{P_2} & \\ 0 & & \boxed{P_r} \end{pmatrix}$$

Chaque bloc P_i étant une matrice markovienne régulière. D'après le théorème 5.2. $(P^d)^n$ tend vers une limite A_0 quand $n \to +\infty$. De plus si $0 \leqslant m < d$ alors P^{nd+m} tend vers $P^m A_0$ quand $n \to +\infty$. La suite P^n possède d sous suites convergentes, elle converge donc au sens de Cesaro vers la moyenne $\frac{1}{d} \sum\limits_{m=o}^{d-1} P^m A_0$ des limites de ses sous-suites. Or P^n converge au sens d'Euler vers A. La limite au sens de Cesaro est donc également A.

6.4.- <u>Exemple</u> (marche aléatoire avec barrières réfléchissantes)

Reprenons l'exemple 4.2.3. où une particule se déplace sur l'ensemble des entiers 1,2,3,4,5 en faisant à chaque instant un pas à droite avec la probabilité p ou un pas à gauche avec la probabilité q. Supposons d'autre part que la particule est "réfléchie" lorsqu'elle arrive en 1 ou 5 et qu'elle retourne d'où elle vient. La matrice des transitions est de la forme :

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Partant d'un état pair la chaîne ne peut se trouver dans un état pair qu'après un nombre pair de pas et dans un état impair qu'après un nombre impair de pas. La chaîne est périodique, de période $d = 2$ et elle comprend deux classes cycliques composées l'une des états pairs, l'autre des états impairs.

Pour déterminer la matrice limite A on résoud l'équation $\vec{\alpha}P = \vec{\alpha}$. Ainsi par exemple si $p = q = \frac{1}{2}$ on trouve $\vec{\alpha} = (\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8})$.

§ 7 - CHAINES ABSORBANTES ET FONCTIONS HARMONIQUES

7.1.- DEFINITION :

1) Soit f une fonction définie sur l'ensemble E des états d'une chaîne de Markov de matrice (p_{ij}). On appelle valeur moyenne de f en $i \in E$, le nombre $\bar{f}(i)$ défini par

$$\bar{f}(i) = \sum_{j \in E} p_{ij} f(j)$$

2) Soit A un sous ensemble de E. On dit que f est harmonique dans A si pour tout $i \in A$ on a $f(i) = \bar{f}(i)$.

3) f est dite surharmonique (resp. sous harmonique) dans A si pour $i \in A$ on a $f(i) \geq \bar{f}(i)$ (resp. $f(i) \leq \bar{f}(i)$)

C'est surtout dans l'étude des chaînes absorbantes que nous avons rencontré des fonctions surharmoniques. Nous allons en donner quelques exemples et voir que cette notion permet de résoudre certaines questions sans calcul matriciel, en se "promenant" sur le graphe d'un point à l'autre.

7.2.- Exemple 1 : Soit R l'ensemble des états absorbants et $\ell \in R$. Soit a la fonction définie sur E par : $a(i)$ = probabilité partant de i d'être absorbé par ℓ. (voir le théorème 4.2.5.). Alors on a :

$$a(\ell) = 1$$

$$a(k) = 0 \quad \text{si} \quad k \in R,\ k \neq \ell$$

$$a(i) = \sum_{j \in E} p_{ij} a(j) \quad \text{si} \quad i \notin R$$

La fonction a est donc harmonique dans $E-R$.

7.3.- Exemple 2 : Dans une chaîne absorbante, soit $\ell \in E-R$ et m la fonction définie sur E par :

$m(i)$ = le nombre moyen de visites en ℓ partant de i.

On a (voir 4.2.3. et 4.2.4.) :

$$m(i) = 0 \quad \text{si} \quad i \in R$$

$$m(\ell) = 1 + \sum_{j \in E} p_{\ell j}\, m(j)$$

$$m(i) = \sum_{j \in E} p_{ij}\, m(j) \quad \text{si} \quad i \notin R,\ i \neq \ell$$

Donc m est surharmonique dans $E-R$ et harmonique dans $E-\{R \cup \{\ell\}\}$.

7.4.- Exemple 3 : Dans une chaîne absorbante, soit n la fonction $i \mapsto n(i)$ = nombre moyen de pas partant de i jusqu'à l'absorbtion. On a :

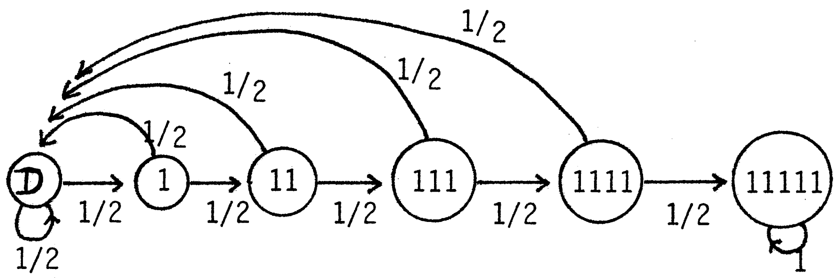
$$n(i) = 0 \quad \text{si} \quad i \in R$$

$$n(i) = 1 + \sum_{j \in E} p_{ij}\, n(j) \quad \text{si} \quad i \notin R$$

Donc n est surharmonique dans $E-R$.

7.5.- Exercice : On joue à pile ou face jusqu'à ce qu'on obtienne une série de cinq succès (pile) consécutifs. Quelle est la durée moyenne du jeu ? (on suppose la pièce régulière).

Solution : Notons 1 pour pile et $\mathcal{D}$ l'état de départ. Dès que face apparaît tout est à recommencer, on revient donc à l'état initial. Ainsi on a une chaîne de Markov :



Renumérotons 0,1,2,3,4,5 les divers états, 5 est l'état absorbant. Notons $n(i)$ le nombre moyen de pas, partant de i, pour atteindre 5. On a $n(5) = 0$. Posons $n(\mathcal{D}) = x$. Alors d'après 7.4., on a :

$$n(4) = 1 + \frac{1}{2}\, n(\mathcal{D}) + \frac{1}{2}\, n(5) = 1 + \frac{1}{2}\, x$$

$$n(3) = 1 + \frac{1}{2}\, x + \frac{1}{2}\left(1 + \frac{1}{2}\, x\right) = \frac{3}{2} + \frac{3}{4}\, x$$

$$n(2) = 1 + \frac{1}{2}\, x + \frac{1}{2}\left(\frac{3}{2} + \frac{3}{4}\, x\right) = \frac{7}{4} + \frac{7}{8}\, x$$

$$n(1) = 1 + \frac{1}{2}\, x + \frac{1}{2}\left(\frac{7}{4} + \frac{7}{8}\, x\right) = \frac{15}{8} + \frac{15}{16}\, x$$

$$n(\mathcal{D}) = \frac{1}{2}\, n(\mathcal{D}) + \frac{1}{2}\, n(1) + 1$$

$$\frac{x}{2} = \frac{15}{16} + \frac{15}{32}\, x + 1 \qquad \text{d'où} \quad x = 62$$

LECTURE LIBRE :

§ 8 - FERMETURE TRANSITIVE D'UN GRAPHE. APPLICATION A LA DETERMINATION DES CLASSES D'ELEMENTS COMMUNICANTS DANS UNE CHAINE DE MARKOV (*)

8.1.- Soit E un ensemble. Un graphe sur E est un couple (E,Γ) où Γ est une application multivoque de E dans E (multivoque signifie qu'un élément de E peut avoir plusieurs images par Γ). Si $y = \Gamma(x)$ nous dirons qu'une flèche part de x vers y.

8.2.- DEFINITION : Soit $G = (E,\Gamma)$ un graphe. La fermeture transitive de Γ est l'application $\hat{\Gamma}$ définie sur E par :

$$\hat{\Gamma}(x) = \{x\} \cup \Gamma(x) \cup \Gamma^2(x) \cup \ldots \cup \Gamma^n(x) \cup \ldots$$

où $\Gamma^2(x) = \Gamma(\Gamma(x))$, $\Gamma^3(x) =$ etc...

8.3.- PROPOSITION : Soit E l'espace des états d'une chaîne de Markov : Soit Γ l'application définie sur E par :

$$\Gamma(x) = \{y \in E / p_{xy} > 0\}$$

p_{xy} = la probabilité de transition de x vers y.
Deux éléments a et $b \in E$ communiquent si et seulement si $b \in \hat{\Gamma}(a)$ et $a \in \hat{\Gamma}(b)$.

Démonstration : évidente.

8.4.- DEFINITION : Matrice indicatrice du graphe : c'est la matrice obtenue à partir de la matrice de transition en remplaçant les $p_{ij} > 0$ par des 1.

Exemple : $$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} \rightsquigarrow M = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Un 1 à la $(i,j)^e$ place indique qu'on peut aller de i vers j en une étape.

(*) ce paragraphe est facultatif : il ne sera pas traité en cours.

8.5.- Multiplication booléienne de 2 matrices :

Les opérations somme et produit Booléiens sont définis par :

$$1+1 = 0+1 = 1+0 = 1 \ ; \ 0+0 = 0 \ ;$$
$$1.1 = 1 \ ; \ 1.0 = 0.1 = 0.0 = 0.$$

La somme booléienne de 2 matrices est définie par la somme booléienne de ses coefficients et la multiplication par :

$$MM' = a_{ij} = \sum_{\text{booléienne}} m_{ik} \underbrace{.}_{\text{produit booléien}} m'_{kj}$$

> *8.6.- PROPOSITION : La matrice M^k (puissance booléienne $k^{\text{ième}}$) est la matrice caractéristique du graphe (E,Γ^k). C'est une matrice dont le coefficient $a_{ij}^{(k)}$ vaut 1 s'il existe un chemin de longueur k allant de i vers j.*

Démonstration : identique à celle définissant la matrice P^k (puissance $k^{\text{ième}}$ de la matrice des transitions).

8.7.- Remarque : La somme booléienne des matrices $I+M+\ldots+M^k$ s'interprête très simplement : le coefficient c_{ij} vaut 1 si $j \in \{i\} \cup \Gamma(i) \cup \Gamma^2(i) \cup \ldots \cup \Gamma^k(i)$ c'est-à-dire s'il existe un chemin de longueur $\leqslant k$ allant de i à j.

> *8.8.- PROPOSITION : Il existe un entier n tel que $\forall k \geqslant n$ la matrice booléienne $I+M+M^2+\ldots+M^k$ est constamment égale à la matrice $I+M+M^2+\ldots+M^n$.*

Démonstration : $\forall i, \exists$ un entier n_i tel que $\hat{\Gamma}(i) = \{i\} \cup \Gamma(i) \cup \ldots \cup \Gamma^{n_i}(i)$ en effet à partir d'un certain moment les $\Gamma^n(i)$ sont inclus dans la réunion des précédants puisque E est un ensemble fini. Posons $n = \sup_i n_i$. Alors ce n convient compte tenu de la remarque.

8.9.- Détermination des classes d'éléments communicants :

On calcule la somme $I+M+M^2+\ldots+M^k = S_k$ jusqu'à ce que $S_k = S_{k+1}$. Alors c'est terminé, toutes les sommes suivantes sont égales. Pour savoir si i et j sont dans la même classe d'éléments communicants il suffit de regarder si on a à la fois $c_{ij} = 1$ et $c_{ji} = 1$ (c_{ij} est le coefficient i,j de la matrice $I+M+\ldots+M^n$).

FIN de la lecture libre.

§ 9 - CHAINES DE MARKOV DENOMBRABLES

9.1.- Dans les précédents paragraphes nous avons obtenu des résultats assez précis dans le cas où l'espace des états E est fini. On va brièvement étudier la classification des états lorsque E est dénombrable. La partition de E en classes d'éléments communicants est toujours valable, mais la notion d'ensembles minimaux doit être abandonnée : pour la remplacer nous allons introduire la notion d'état récurrent. On supposera les états de E numérotés, et on prendra ainsi $E = \mathbb{N}$. On pourra continuer à considérer les p_{ij} comme formant une matrice (à une infinité de lignes et de colonnes). On notera $p_{ij}^{(n)}$ la probabilité d'aller de i en j en n étapes. La matrice des $p_{ij}^{(n)}$ est la matrice P^n. Supposons que la chaîne parte de i et soit α_n = la probabilité que la chaîne retourne en i pour la première fois à l'instant n. Alors $\alpha = \sum_{n=1}^{\infty} \alpha_n$ est la probabilité que la chaîne retourne en i (tôt ou tard).

9.2.- DEFINITION : *On dit que l'état i est récurrent si $\alpha = 1$ et i est transient si $\alpha < 1$.*

9.3.- PROPOSITION : *i est récurrent si et seulement si $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$*

Démonstration : Posons $\beta_n = p_{ii}^{(n)}$. Alors on a la relation

$$\beta_n = \beta_0\alpha_n+\beta_1\alpha_{n-1}+\dots+\beta_{n-1}\alpha_1+\beta_n\alpha_0 \qquad (\beta_0=1,\ \alpha_0=0) \quad n=1,2,\dots$$

(formule de la probabilité totale appliquée avec le système complet d'événements A_k = la chaîne retourne en i pour la première fois après k pas $(k = 0,\dots,n)$).

Considérons les séries formelles $\beta(x) = \sum_{n=0}^{\infty} \beta_n x^n$ et $\alpha(x) = \sum_{n=0}^{\infty} \alpha_n x^n$

La relation de récurrence peut s'écrire $\beta(x)-1 = \beta(x)\alpha(x)$. Comme $\alpha(x)$ admet un rayon de convergence égal à 1, il en est de même pour $\beta(x)$. Ainsi $\beta(x) = \dfrac{1}{1-\alpha(x)}$.

Or i est récurrent si et seulement si $\alpha = \sum_{n=0}^{\infty} \alpha_n = \lim_{x\to 1} \alpha(x) = 1$ ie : $\lim_{x\to 1} \beta(x) = \infty$, ie : $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

9.4.- THEOREME :

1) *Si i est récurrent alors il y a une probabilité égale à un pour que la chaîne retourne une infinité de fois en i.*

2) *Si i est transient il y a une probabilité 1 pour que la chaîne ne passe en i qu'un nombre fini de fois.*

La démonstration nécessite un lemme important :

9.4.1.- LEMME *(de Borel-Cantelli)*

Soit A_n une suite d'événements de probabilités respectives $p_n = P(A_n)$ $(n = 0,1,\dots)$.

Si $\sum_{n=0}^{\infty} p_n < +\infty$, alors avec une probabilité égale à 1 il y a seulement un nombre fini d'événements A_n qui peuvent être réalisés simultanément.

Démonstration : Soit B l'événement : "seuls un nombre fini d'événements A_n peuvent se réaliser simultanément" et soit B_n l'événement : "au moins un des A_k est réalisé pour $k > n$".

On a évidemment $B_n = \bigcup_{k=n}^{\infty} A_k$ et $\bar{B} = \bigcap_{n=0}^{\infty} B_n$.

De plus comme $B_1 \supset B_2 \supset B_3 \supset \dots$ la suite B_n est décroissante et d'après la continuité de P par suites décroissantes (voir 2.6 chapitre 2). On a :

$$P(\bar{B}) = P(\bigcap_{n=0}^{\infty} B_n) = \lim_{n\to\infty} P(B_n)$$

Or $\quad P(B_n) \leqslant \sum_{k=n}^{\infty} P(A_k) = \sum_{k=n}^{\infty} p_k \to 0$ comme reste d'une série convergente. On a donc $P(\bar{B}) = 0$ ie : $P(B) = 1$, cqfd.

Démonstration du théorème : Soit T_1 l'instant du premier retour de la chaîne en i, de même notons T_k l'instant du k^e retour en i. On a $P(T_1 < \infty) = \alpha$ et il est clair que $P(T_k < \infty | T_{k-1} < \infty) = \alpha$ donc $P(T_k < \infty) = \alpha^k$.

Si i est transient on a $\alpha < 1$ donc

$$\sum_{k=1}^{\infty} P(T_k < \infty) = \sum_{k=1}^{\infty} \alpha^k < \infty$$

et d'après le lemme seul un nombre fini d'événements $(T_k < \infty)$ peuvent être réalisés d'où la partie 2) du théorème.

Si i est récurrent, $\alpha = 1$ et pour tout k, $P(T_k < \infty) = 1$. Désignons par N le nombre de retours en i. On a évidemment

$$[N \geqslant k] = [T_k < \infty]$$

et $$[N = \infty] = \bigcap_k [T_k < \infty]$$

Comme P est continue on a

$$P(N = \infty) = \lim_{k\to\infty} P(T_k < \infty) = 1$$

D'où l'assertion 1) du théorème.

9.5.- COROLLAIRE :

1) Tout état accessible à partir d'un état récurrent est récurrent.

2) Les états d'une même classe d'éléments communicants sont tous de même nature (ie : tous récurrents ou tous non récurrents).

Démonstration :

1) Soit i un état récurrent et soit j un état tel que $i \to j$ (ie : il existe un entier m tel que $p_{ij}^{(m)} > 0$) alors on a $j \to i$ (i est accessible à partir de j) car sinon la probabilité de retour en i serait $\leqslant 1-p_{ij}^{(m)}$ donc i ne serait pas récurrent. (En effet : "atteindre j en m étapes"" $\Rightarrow$ "atteindre j" $\Rightarrow$ "non retour en i" donc P(retour en i) + $p_{ij}^{(m)} \leqslant 1$). Il reste à montrer que j est récurrent mais ceci va résulter de 2).

2) Soient i et j communicants. Il existe des entiers n et m tels que $p_{ij}^{(n)} = a > 0$ et $p_{ji}^{(m)} = b > 0$. Alors pour k assez grand $(k>n+m)$ on a :

$$p_{ii}^{(k)} \geqslant p_{ij}^{(n)} \; p_{jj}^{(k-n-m)} \; p_{ji}^{(m)} = ab p_{jj}^{(k-n-m)}$$

et de même

$$p_{jj}^{(k)} \geqslant ab \; p_{ii}^{(k-n-m)}$$

Il en résulte que les séries $\sum_{k=o}^{\infty} p_{ii}^{(k)}$ et $\sum_{k=o}^{\infty} p_{jj}^{(k)}$ sont de même nature (convergentes ou divergentes).

9.6.- Exercice : Une particule effectue une marche aléatoire sur l'ensemble $\mathbb{Z}$ des entiers relatifs. A chaque instant elle se déplace d'une unité à gauche avec la probabilité p ou d'une unité à droite avec la probabilité $q = 1-p$. Montrer que si $p = q = \frac{1}{2}$ tous les états sont récurrents et si $p \neq q$ avec $4pq < 1$ tous les états sont transients.

Solution :

On a $$p_{ii}^{(2n)} = C_{2n}^{n} \; p^n q^n = \frac{(2n)!}{n!n!} \; p^n q^n \sim \frac{(4pq)^n}{\sqrt{\pi n}}$$

(d'après la formule de Stirling : $n! \sim \sqrt{2\pi n} \; n^n e^{-n}$).

1) Si $p = q = \frac{1}{2}$ alors $p_{ii}^{(2n)} \sim \frac{1}{\sqrt{\pi n}}$ donc $\Sigma p_{ii}^{(n)} = +\infty$ et la chaîne est récurrente.

2) Si $4pq < 1$ la série $\sum_{n} \frac{(4pq)^n}{\sqrt{\pi n}}$ converge donc les états sont non récurrents.

9.7.- Exercice : Soit une marche aléatoire sur $\mathbb{Z}$ (voir exercice 8.6.) avec $p = q = \frac{1}{2}$ (donc tous les états sont récurrents). Soit T_1 le temps du premier retour en 0. Montrer que $E(T_1) = \infty$.

Solution : Avec les notations de .3 soit α_n = la probabilité de retour en i pour la première fois à l'instant n. La fonction génératrice de T_1 est $\alpha(x) = \sum_{n=o}^{\infty} \alpha_n x^n$. On sait d'après la démonstration de 8.3. que l'on a :

$$\beta(x) = \sum_{n=o}^{\infty} \beta_n x^n = \frac{1}{1-\alpha(x)}$$

D'après 8.6 on a $\beta_{2n} = C_{2n}^{n} \; (pq)^n$, $\beta_{2n+1} = 0$.

Donc $\beta(x) = \dfrac{1}{\sqrt{1-4pqx^2}}$

D'où $\alpha(x) = 1-\sqrt{1-4pqx^2} = 1-\sqrt{1-x^2} \quad (p = q = \frac{1}{2})$.

Ainsi $E(T_1) = \alpha'(1) = \infty$.

9.8.- Remarque : On vient de voir dans l'exercice précédent que bien que 0 soit récurrent (ie : on y revient une infinité de fois) le temps moyen de retour en 0 est infini ! Voilà qui semble paradoxal mais qui illustre bien les difficultés du cas où E est dénombrable.

On peut réaliser concrètement la situation de .7 par une partie illimitée de pile ou face (+1 pour pile, -1 pour face).

On retourne une infinité de fois à l'équilibre (autant de piles que de faces) mais le temps moyen du premier retour à l'équilibre est infini !