REVUE DE STATISTIQUE APPLIQUÉE

B. LE MAIRE
G. MAUFFREY

Chaînes de Markov, non homogènes, à causalité constante : matrices d'écart d'homogénéité

Revue de statistique appliquée, tome 23, nº 4 (1975), p. 41-76
<http://www.numdam.org/item?id=RSA_1975__23_4_41_0>

© Société française de statistique, 1975, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (http://www.sfds.asso.fr/publicat/rsa.htm) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

CHAINES DE MARKOV, NON HOMOGÈNES, à CAUSALITÉ CONSTANTE : MATRICES D'ÉCART D'HOMOGÉNÉITÉ (*)

B. Le MAIRE et G. MAUFFREY
Centre d'Enseignement Supérieur des Affaires Jouy-en-Josas

SOMMAIRE

Cet article est issu d'une recherche sur le développement de certains modèles stochastiques dans le domaine du *comportement du consommateur,* vu sous l'angle *Marketing.*

Le modèle général que nous présentons ici est une extension, que nous croyons à la fois réaliste et fructueuse, de modèles *bien connus* de *"Brand Loyalty and Brand Switching"*, les modèles *Markoviens, homogènes, du 1er ordre.*

(Ces "anciens" modèles, bien qu'assez rudimentaires et assez simplistes continuent d'être employés dans la pratique, en vue de donner une première approximation quantitative de *l'évolution* au cours du temps de l'état d'un système qui peut être, par exemple, la part de marché d'une marque d'un produit particulier (eau minérale, biscuit sec, bière, . . .) qui appartient le plus souvent à la classe des produits alimentaires de grande consommation).

L'intérêt d'un modèle un peu plus sophistiqué apparaît d'autant mieux que l'on considère avec plus d'acuité les inconvénients du modèle Markovien "habituel" ; de fait, le plus gros reproche que l'on peut faire à ce modèle classique, mis sur pied sous sa forme actuelle aux alentours des années 1960, a été fait par Ehrenberg(1), qui déclarait en substance : *Les analystes "Markoviens", pour justifier leurs hypothèses d'homogénéité (au cours du temps) du phénomène étudié,* considèrent que la situation analysée est dans une *phase stable,* (marché à l'équilibre, ou à peu près) et qu'il n'y a donc pas de raison que la *"matrice de transition*(2) ne soit pas, elle aussi, *stable,* c'est-à-dire, *"indépendante du temps".*

Ehrenberg concluait alors "si rien ne bouge, ou à peu prés, même si l'hypothèse d'homogénéité est justifiée, l'intérêt d'une telle analyse est presque nul".

Il est évidemment plus intéressant de pouvoir étudier une situation moins figée, et c'est ce que nous allons faire dans cet article.

Cet article se décompose, de façon schématique, en 2 parties.

---

(1) An appraisal of Markov Brand-Switching Models – Journal of Marketing Research – Vol. 11 (novembre 1965), 437-62.

(2) Cette matrice est aussi appelée tableau ou matrice de comportement.

(*) Article remis en Avril 1974, révisé en Septembre 1974.

La 1ère partie établit les résultats *purement théoriques* concernant ce modèle, qui est une extension *évolutive, (non homogène dans le temps)* des modèles Markoviens usuels. Nous *démontrons* un certain nombre de résultats nouveaux, nous *précisons* certaines conjectures (dont celle d'Harary-Lipsten) en les généralisant à un nombre quelconque d'états et en donnant certains *contre-exemples,* non mis en lumière jusqu'alors, résultant du *relâchement* de certaines hypothèses, en utilisant pour cela des résultats relativement classiques d'algèbre linéaire. Nous prouvons en particulier que certains résultats asymptotiques de chaînes homogènes ergodiques restent valables.

La 2ème partie, beaucoup plus *orientée vers l'utilisation* du modèle expose un certain nombre d'algorithmes indispensables pour l'utilisation effective de notre modèle. Ces algorithmes développent essentiellement une méthodologie de recherche de la matrice de causalité dans le cadre d'une chaîne à 2 états. Ces algorithmes vont du plus général (connaissance de certains vecteurs d'états seulement) au plus particulier (connaissance des matrices de transition).

Le problème se pose en effet, pratiquement de la façon suivante : quelles données sont disponibles pour évaluer et tester les paramètres du modèle.

1/ Si les parts de marché sont seules disponibles, cas assez fréquent hélas, l'algorithme correspondant doit être assez complexe analytiquement, et débouche sur des problèmes statistiques relativement sophistiqués si l'on veut tenir compte des *fluctuations aléatoires* (non expliquées par le modèle). Nous avons mis au point dans ce cas un algorithme analytique et construit 5 ou 6 programmes empiriques informatiques dans ce domaine.

2/ Si, au contraire, les matrices de transition sont disponibles, cas peu fréquent en pratique (mais considéré souvent comme une hypothèse de base) l'algorithme analytique prend une forme très élémentaire, et les programmes statistiques, mis aussi au point par ailleurs, ne sont qu'une forme à peine élaborée des programmes usuels de régression multiple. Dans cette même partie nous parlerons des problèmes laissés ouverts par cette recherche et des débouchés pratiques que peut présenter un tel modèle.

1ère PARTIE – RESULTATS THEORIQUES

Dans cette première partie, théorique, nous allons introduire la notion générale des chaînes à *causalité constante.* Nous démontrerons en particulier qu'une partie de la conjecture de Lipstein[3] sur les problèmes de convergence d'une telle suite est valide : nous l'avons d'ailleurs généralisée au cas d'une chaîne à un nombre *quelconque* d'états (voir théorème 1, paragraphe III,

---

(3) Lipstein : "A Mathematical Model of Consumer Behavior" *J. of M.*R/2/259-265 (1965)

1ère partie). Nous démontrerons par ailleurs que, dans certains cas, les propriétés limites des chaînes homogènes sont conservées, c'est-à-dire que le produit $M_1\ CM_1\ C^2\ M_1 \ldots C^n\ M_1$ a la même limite que le terme $C^n M_1$, lorsque $n \to \infty$, ce qui n'avait pu être fait par Harary-Lipstein[4].

Soit $\mathcal{C}$ une chaîne de Markov (à nombre fini N d'états) ayant pour suite de vecteurs d'états la suite $(P_n)_{n \in N^*}$ et pour suite de matrices de transition, à une étape, la suite $(M_n)_{n \in N^*}$.

Nous supposerons, dans toute la suite, que la suite (M_n) n'est pas nécessairement constante, c'est-à-dire que la chaîne n'est pas nécessairement homogène, mais que l'on a du moins l'hypothèse restrictive suivante :

Hypothèse : $\exists C$, matrice carrée d'ordre N, $\forall\ n \in N^*$, $M_{n+1} = C\ M_n$, c'est-à-dire qu'il existe une matrice C, dite de *"causalité" markovienne,* telle que la matrice de transition pour la période n + 1 se déduit de la matrice de transition pour la période n par prémultiplication de M_n par C.

Définition : Nous dirons que nous avons *une chaîne à causalité constante.* (Le cas C = I nous ramenant au cas particulier des chaînes homogènes).

Remarque : S'il existe au moins une matrice M_n inversible, C sera obtenue[5] immédiatement de M_n et de M_{n+1}, puisque : $C = M_{n+1} \cdot M_n^{-1}$;

I – QUELQUES DEFINITIONS ET PROPRIETES GENERALES

Définition 1 : Une matrice *pseudo-stochastique* est une matrice dont la somme de *chaque ligne* est égale à 1[6].

Lemme :

i) L'ensemble des *matrices carrées pseudo-stochastiques* (de même ordre) est un *demi-groupe*[7] possédant un élément neutre, la matrice unité

$$\begin{bmatrix} 1 & & 0 \\ & 1\ddots & \\ 0 & & 1 \end{bmatrix}$$

ii) Le sous-ensemble des matrices *carrées pseudo-stochastiques inversibles* est un groupe[8].

- - - - - - - - - - - - - - -

(4) Harary-Lipstein-Styan : "A Matrix Approach to Non-Stationary Chaîns" (Op. Res/ 18/1168-1181/1970).

(5) Nous verrons dans la 2ème partie comment C peut être obtenue pratiquement à partir de données réelles.

(6) Une matrice stochastique étant pseudo-stochastique et vérifiant, en plus, la propriété que tous ses termes soient $\geqslant 0$.

(7) Pour la multiplication matricielle.

(8) Une matrice stochastique étant pseudo-stochastique et vérifiant, en plus, la propriété que tous ses termes soient $\geqslant 0$.

iii) L'ensemble des *matrices carrées pseudo-stochastiques* est un ensemble *convexe.*

Les démonstrations, élémentaires, ne seront pas données ici.

II – ETUDE D'UN CAS PARTICULIER : N (nombre d'états) = 2

A – Quelques notations : (La chaîne considérée, à 2 états, est supposée à causalité constante).

Posons

$$\begin{cases} M_1 = M_2 + \begin{bmatrix} -\epsilon_1 + \epsilon_1 \\ -\epsilon_2 + \epsilon_2 \end{bmatrix} = C^{-1} M_2 & \text{(si on suppose que C est } \textit{inversible}^{(9)} \\ \qquad\qquad\qquad\qquad = DM_2 , & \text{en posant } D = C^{-1} \\ M_2 = M_3 + \begin{bmatrix} -\epsilon_3 + \epsilon_3 \\ -\epsilon_4 + \epsilon_4 \end{bmatrix} = DM_3 & \\ D = \begin{bmatrix} d_1 & 1 - d_1 \\ d_2 & 1 - d_2 \end{bmatrix} & \text{(D est en effet pseudo-stochastique)} \end{cases}$$

Nous allons maintenant relier ϵ_i , i = 1 , 2 , 3 , 4 à d_i , i = 1 , 2

B – Théorème : Si $d_1 \neq d_2$ (voir note en bas de page pour le cas contraire[10],

alors :

$$\begin{bmatrix} \epsilon_3 - \epsilon_3 \\ \epsilon_4 - \epsilon_4 \end{bmatrix} = \frac{1}{d_1 - d_2} \begin{bmatrix} \epsilon_1 - \epsilon_1 \\ \epsilon_2 - \epsilon_2 \end{bmatrix} = (c_1 - c_2) \begin{bmatrix} \epsilon_1 - \epsilon_1 \\ \epsilon_2 - \epsilon_2 \end{bmatrix}$$

---

(9) Cette restriction n'est pas très forte et, de plus, correspond à un cas particulier très facile à étudier directement : en effet, dans le cas N = 2 C non inversible signifie, puisque C est pseudo-stochastique, que les 2 lignes de C sont identiques. On peut alors voir, immédiatement, que l'on a $M_{n+1} = M_n$, dès que $n \geqslant 1$, avec, de *plus,* les 2 lignes de M_n identiques.

On a alors, toujours pour $n \geqslant 1$, non seulement un processus homogène, mais aussi un processus stationnaire.

(10) Si on pose $k = d_1 - d_2$ on peut voir immédiatement, de la démonstration du théorème que

$$k = \frac{m_1 - p_1}{m_2 - p_2}$$

dans le cas où M_2 est *inversible.* Si M_2 n'est pas inversible, la chaîne est *homogène et stationnaire.*

Démonstration du théorème[11]

1/ Posons

$$M_2 = \begin{bmatrix} m_2 & 1 - m_2 \\ p_2 & 1 - p_2 \end{bmatrix} \qquad \left[M_1 = \begin{bmatrix} m_1 & 1 - m_1 \\ p_1 & 1 - p_1 \end{bmatrix} \right]$$

La relation $D \times M_2 = M_2 + \begin{bmatrix} -\epsilon_1 + \epsilon_1 \\ -\epsilon_2 + \epsilon_2 \end{bmatrix}$ $(= M_1)$ donne immédiatement, par identification :

$$d_1 m_2 + (1 - d_1)\, p_2 = m_2 - \epsilon_1$$
$$d_2 m_2 + (1 - d_2)\, p_2 = p_2 - \epsilon_2$$

c'est-à-dire :

$$d_1 (m_2 - p_2) = m_2 - p_2 - \epsilon_1$$
$$d_2 (m_2 - p_2) = - \epsilon_2$$

et, en supposant M_2 *inversible,*

$$(E_1) \quad \boxed{\begin{array}{lll} d_1 = 1 - \dfrac{\epsilon_1}{m_2 - p_2} & \text{ou} & \epsilon_1 = (m_2 - p_2)(1 - d_1) \\[2ex] d_2 = \dfrac{-\epsilon_2}{m_2 - p_2} & \text{ou} & \epsilon_2 = (p_2 - m_2)\, d_2 \end{array}}$$

2/ Si M_2 *n'est pas inversible,* c'est-à-dire si $m_2 = p_2$, on a alors, nécessairement :

$$M_1 = \begin{bmatrix} d_1 & 1 - d_1 \\ d_2 & 1 - d_2 \end{bmatrix} \qquad \begin{bmatrix} m_2 & 1 - m_2 \\ m_2 & 1 - m_2 \end{bmatrix}$$
$$= \begin{bmatrix} m_2 & 1 - m_2 \\ m_2 & 1 - m_2 \end{bmatrix} = M_2$$

Le processus est alors *homogène et stationnaire.*

- - - - - - - - - - - - - -

(11) Une démonstration plus rapide sera donnée dans la partie E de ce même paragraphe, mais les liens entre les d_i, les c_j et les ϵ_k ont paru intéressants à étudier directement.

3/ De même, en utilisant la relation :

$$C.M_2 = M_2 + \begin{bmatrix} \epsilon_3 & -\epsilon_3 \\ \epsilon_4 & -\epsilon_4 \end{bmatrix} \qquad (= M_3)$$

on obtient (pour $m_2 \neq p_2$)

$$(E_2) \quad \boxed{\begin{array}{lll} c_1 = 1 + \dfrac{\epsilon_3}{m_2 - p_2} & \text{ou} & \epsilon_3 = (1 - c_1)(p_2 - m_2) \\ c_2 = \dfrac{\epsilon_4}{m_2 - p_2} & \text{ou} & \epsilon_4 = c_4 (m_2 - p_2) \end{array}} \quad (12)$$

4/ Les relations

$$(E_3) \quad \boxed{c_1 = \frac{1 - d_2}{d_1 - d_2} \quad \text{et} \quad c_2 = \frac{-d_2}{d_1 - d_2}}$$

liant les coefficients des matrices inverses C et D, jointes aux relations (E_1) et (E_2) donnent alors :

$$\boxed{\epsilon_3 = \frac{\epsilon_1}{d_1 - d_2} \quad \text{et} \quad \epsilon_4 = \frac{\epsilon_2}{d_1 - d_2}}$$

c.q.f.d.

C – Corrolaire :

i) *Si* $\alpha_1 \neq \alpha_2$ (M_2 inversible) alors

1/ si $\boxed{0 < |c_1 - c_2| < 1}$

$$M_n \to M_1 + \frac{1}{1 - (c_1 - e_2)} \begin{bmatrix} +\epsilon_1 & -\epsilon_1 \\ +\epsilon_2 & -\epsilon_2 \end{bmatrix}$$

2/ si $\boxed{c_1 - c_2 | > 1}$ $\lim M_n$ n'existe pas (13)

---

(12) Rappelons que l'on peut montrer aisément que :

$$c_1 - c_2 = \frac{m_2 - p_2}{m_1 - p_1}$$

(13) Ce cas ne peut d'ailleurs se produire dans la pratique, car pour n suffisamment grand, M_n ne serait plus stochastique.

3/ si $\boxed{c_1 - c_2 = 1}$ $\rightarrow$ (*)

$$\begin{cases} 1 \quad c_1 = 1 \quad M_n = M_1 \ , \forall\, n \in N^* \\ 2 \quad c_1 \neq 1 \quad \lim M_n \text{ n'existe pas} \end{cases}$$

4/ si $\boxed{c_1 - c_2 = 1}$

$$\begin{cases} M_{2n+1} = M_1 \ , \forall\, n \in N \\ M_{2n+2} = M_2 \ , \forall\, n \in N \end{cases}$$

5/ si $\boxed{c_1 - c_2 = 0}$, le processus est homogène, et stationnaire

ii) Si $\alpha_1 = \alpha_2$ $M_n = M_2$, $\forall\, n \geqslant 2$ (Le processus est homogène, et stationnaire).

Démonstration : La preuve du corrolaire est immédiate, puisque reposant sur les propriétés bien connues de la série géométrique.

D – Stochasticité de la limite (éventuelle) de la suite $(M_n)_{n \in N}$

Nous allons maintenant nous intéresser, dans le cas où (M_n) converge, aux propriétés nécessaires de cette limite.

Les propriétés des matrices de transition sont telles que l'on doit avoir, $\forall\, n \in N$, M_n stochastique. On a d'abord le lemme suivant :

Lemme : *Si lim M_n existe, alors cette limite est stochastique.*

Preuve : La limite éventuelle de cette suite doit aussi être stochastique, l'ensemble des matrices stochastiques étant un sous-ensemble convexe borné fermé.

Remarque : Nous avons vu, dans le paragraphe C, dans quelles conditions la suite $(M_n)_{n \in N}$, c'est-à-dire en fait la suite

$$(C^{n-1} \, . \, M_1)$$

convergeait, au moins dans le cas N = 2, N étant le nombre d'états.

Nous allons généraliser ce résultat dans le cas N quelconque, en montrant que si toutes les valeurs propres de C, en dehors d'une seule (qui est toujours égale à 1), sont *strictement inférieures, en module, à 1,* alors la suite (M_n) converge : de plus, quelle que soit la matrice C, si la suite (M_n) converge, c'est nécessairement vers une matrice stochastique dans le cas où toutes les M_n sont stochastiques.

Avant d'établir ce résultat, très général, nous allons donner une autre démonstration du théorème B.

E – Autre forme du théorème B

Théorème : Si E_1 , E_2 , . . . , E_n sont *les différentes matrices d'écart,*

alors : $E_n = \lambda^{n-1} E_1$, λ et 1[14]

- - - - - - - - - - - - - -

(14) C a toujours 1 comme valeur propre, puisqu'elle est pseudo-stochastique.

étant les *2 valeurs propres* de la matrice C, avec $\lambda = \frac{1}{k} = c_1 - c_2$

(Rappels : $M_2 = C \cdot M_1 = M_1 + E_1, \ldots, M_{n+1} = C \cdot M_n + E_n$)

Preuve : D'après le théorème de Cayley, nous avons :

$$(C - I)(C - \lambda I) = 0$$

d'où, en post multipliant par M_1, première matrice de transition

$$(C - I)\, CM_1 = \lambda (C - I) M_1$$

$$\Longleftrightarrow (C - I)\, M_2 = \lambda (C - I) M_1$$

$$\Longleftrightarrow \boxed{E_2 = \lambda E_1}$$

C^n ayant les mêmes valeurs propres que C, nous aurions, de même :

$$\boxed{E^{n+1} = \lambda^n E}$$

Puisque $\begin{bmatrix} c_1 & 1 - c_1 \\ c_2 & 1 - c_2 \end{bmatrix}$ a pour valeurs propres 1 et λ, on a, par identification :

$$\boxed{c_1 - c_2 = \lambda = \frac{1}{k}} = \frac{m_2 - p_2}{m_1 - p_1}$$

c.q.f.d.

III – EXISTENCE D'UNE MATRICE LIMITE DANS LE CAS GENERAL D'UNE CHAINE CAUSALE A N ETATS

A – Position du problème *(conjecture de LIPSTEIN)*

Soit C une matrice de causalité *régulière* d'ordre N. Soit M_1 une matrice carrée *stochastique* d'ordre N. Nous nous proposons de trouver une condition *nécessaire et suffisante* pour que la suite $(M_n)_{n \geqslant 1}$ soit *convergente,* avec

$$\boxed{M_n = C^{n-1} M_1}$$

Nous aurons, pour cela, besoin de quelques propriétés des polynomes *minimaux* d'une matrice.

B – Polynome minimal d'une matrice

Etant donné une matrice A il *existe un polynome* P(X), *normé, de degré minimum* tel que P(A) = 0. Ce polynome, dit *polynome minimal,* divise,

en particulier, le polynome caractéristique de A. On a, plus précisément :

$$P(X) = (X - z_0)^{\beta_0} (X - z_1)^{\beta} I \ldots (X - z_k)^{\beta_k}$$

avec $\begin{cases} z_0, z_1, \ldots, z_k \text{ valeurs propres, } z_i \text{ de } \textit{multiplicité } \alpha_i, \textit{d'indice } \beta_i \\ \text{et} \quad i = 0, 1, \ldots k, 1 \leqslant \beta_i \leqslant \alpha_i, \sum_{i=0}^{k} \alpha_i = N \end{cases}$

Définition : z_i est valeur propre à *spectre simple* ssi la dimension du sous-espace vectoriel propre associé à z_i est α_i. On peut démontrer que :

Théorème : z_i est à *spectre simple* ssi z_i est *d'indice 1* (i.e l'exposant de $(X - z_i)$ dans le polynome *minimal* est égal à 1)

Remarque : Si *toutes* les valeurs propres d'une matrice d'ordre N sont à *spectre simple,* la matrice est dite *scindée.* Cela signifie que R^N est somme directe des sous-espaces *propres,* qui sont, ici égaux aux sous-espaces *spectraux.*

C – Propriétés générales des matrices d'écart

Rappelons qu'une matrice *d'écart* E est une matrice dont la somme des éléments de chaque ligne vaut zéro.

La chaîne de Markov, à N états, étudiée ici étant supposée à causalité constante, nous avons :

$$M_{n+1} = C^n M_1$$

et nous *poserons :*

$$\boxed{E_{n+1} = M_{n+2} - M_{n+1} = (C - I)\, C^n M_1}$$

Soit : $X^r + A_{r-1} X^{r-1} + \ldots + A_0$ le polynome minimal de C.

1 *étant valeur propre de* C, nous pouvons écrire :

$$\begin{cases} A_{r-1} = -1 + B_{r-2} \\ A_{r-2} = -B_{r-2} + B_{r-3} \\ \vdots \\ A_p = -B_p + B_{p-1} \\ \vdots \\ A_1 = -B_1 + B_0 \\ A_0 = -B_0 \end{cases}$$

L'équation minimale peut donc s'écrire en faisant intervenir C :

$$\boxed{(C - I)C^{r-1} + B_{r-2}(C - I)C^{r-2} + \ldots + B_0(C - I) = 0} \quad \text{(Eq4)}$$

D'où :

Lemme 1 : *Les matrices d'écart* E_n *vérifient une équation de récurrence linéaire à coefficients constants, d'ordre* $r - 1$, r *étant le degré du polynome minimal de* C.

En effet l'équation (E_n) devient, en la multipliant par $C^n M_1$:

$$(C - I)C^{r+n-1}M_1 + B_{r-2}(C - I)C^{r+n-2} + \ldots + B_0(C - I)C^n M_1 = 0$$

et, en remplaçant $(C - I)C^n M_1$ par E_{n+1}

$$\boxed{E_{n+r} + B_{r-2}E_{n+r-1} + \ldots + B_0 E_{n+1} = 0} \quad \text{(Eq5)}$$

c.q.f.d

Lemme 2 : *Quelque soit* $n \geqslant 1$, E_n *peut s'écrire comme combinaison linéaire des* $(r - 1)$ *premières matrice d'écart*

$$E_n = \sum_{i=1}^{r-1} a(i, n) E_i$$

les $a(i, n)$ *étant déterminés de façon unique par des équations de récurrence linéaire à coefficients constants, lorsque les* E_i, $i = 1$ *à* $r - 1$ *sont indépendants.*

Démonstration :

i) Par récurrence il est évident que les E_n peuvent s'exprimer comme *combinaisons linéaires* des $r - 1$ premières matrices d'écart, c.a.d :

$$E_n = \sum_{i=1}^{r-1} a(i, n) E_i$$

ii) Par identification[15], les $a(i, n)$ doivent donc vérifier les équations :

$$\boxed{1 \leqslant i \leqslant r - 1,\ a(i, n + r - 1) + B_{r-2}\, a(i, n + r - 2) + \ldots + B_0\, a(i, n) = 0} \quad \text{(Eq6)}$$

Remarque : Cette équation de récurrence se résoud de façon *simple.*

En effet : B_{r-2} représente *l'opposé* de la somme des $r - 1$ valeurs spectrales de C (différentes ou non), 1 étant exclu (1 fois).

B_{r-3} représente le produit 2 à 2 des $r - 1$ valeurs spectrales.

B_{r-4} représente l'opposé du produit 3 à 3.

. . .

- - - - - - - - - - - - - - -

(15) Cette identification étant possible, puisque les $(E_i)_{i=1, \ldots r-1}$ sont supposées indépendantes.

Ces racines (spectrales) étant notées z_i, $0 \leqslant i \leqslant k$, z_i étant d'indice β_i, avec $\sum_{i=0}^{k} \beta_i = r - 1$, il est bien connu que les solutions de l'équation de de récurrence (Eq6) sont de la forme :

$$\boxed{a(p, n) = \sum_{i=0}^{k} \sum_{j=1}^{\beta_i} \lambda(i, j, p) n^{j-1} z_i^n \quad \text{pour} \quad 1 \leqslant p \leqslant r - 1} \qquad \text{(Eq7)}$$

les $\lambda(i, j, p)$ *étant déterminés* par les conditions initiales :

$$(a(p, k) = \delta_{pk} \quad \text{pour} \quad 1 \leqslant k \leqslant r - 1)$$

Lemme 3 : $\forall i = 0, 1, \ldots k \; \forall j = 1, \ldots, \beta_i \; \exists \; p$ t.q. $\lambda(i, j, p) \neq 0$

En effet, dans $\mathbf{R}^{r-1}$ les vecteurs $e_p = (\delta_{p\ell})_{\ell = 1, \ldots r-1}$ représentent la base canonique : si le lemme était faux cela signifierait que ces $r - 1$ vecteurs indépendants seraient chacun combinaison linéaire de $r - 2$ vecteurs, ce qui est absurde c.q.f.d

On a donc, en particulier, $\forall i \; \exists j \; \exists p$ t.q. $\lambda(i, j, p) \neq 0$

On a donc le

Théorème 1 : *La suite*

$$(M_n)_{n \geqslant 1}, \quad \text{avec} \quad M_{n+1} = M_1 + \sum_{i=1}^{n} E_i \, (= C^n M_1)$$

est convergente, dans le cas où les $r - 1$ *premières matrices* E_n *sont indépendantes, ssi les conditions suivantes sont satisfaites.*

C_1 : 1 *est valeur propre à spectre simple*

C_2 : *les valeurs propres de* C *différentes de* 1 *sont de module strictement inférieur à* 1.

Démonstration

On sait en effet que la série de terme général

$$u_n = n^a z_i^n \quad \text{est convergente} \quad \text{ssi } |z_i| < 1 \quad (\text{avec } a \geqslant 0)$$

Du lemme 2, de l'équation (Eq7) et du lemme 3 on déduit alors le théorème précédent, puisque :

$$E_n = \sum_{p=1}^{r=1} \left(\sum_{i=0}^{k} \sum_{j=1}^{\beta_i} \lambda(i, j, p) \, n^{j-1} \right) E_p \qquad \text{(Eq7bis)}$$

Ce théorème correspond à la conjecture de Lipstein reprécisée dans l'article de Harary-Lipstein-Styan "a matrix approach to nonstationary chains" (Op. Res. N-D 1970) dont la démonstration et les conditions de suffisance n'avaient pas encore été énoncées, à notre connaissance.

Corollaire :

Si 1 *est l'unique valeur propre de* C, *alors la suite* $(M_n)_{n \geqslant 1}$ *est convergente ssi* $C = I$

On a aussi le :

Théorème 2 : *La suite* (M_n) *est convergente au sens de CESARO dans le cas des* $(E_n)_{n=1,\ldots,r-1}$ *indépendantes ssi les conditions suivantes sont vérifiées.*

C_{1bis} : Toutes les valeurs propres de module 1 sont à spectre simple.

C_{2bis} : Les autres valeurs propres sont de module strictement inférieur à 1.

IV – RECHERCHE DE LA FORME DE LA MATRICE LIMITE DANS LE CADRE DU THEOREME, C ETANT UNE MATRICE SCINDEE

A – Notations

Soit $z_1, z_2, \ldots, z_{r-1}$ les valeurs spectrales de C autres que 1.

Nous poserons :

$$\ell_i(z) = \prod_{j \neq i} \frac{z - z_j}{z_i - z_j} \quad \text{avec} \quad (\ell_i(z_j) = \delta_{ij})$$

B – Expression des λ

L'équation Eq 7 peut ici s'écrire, puisque $\beta_i = 1 \forall i$ (donc $j = 1$) en posant $\lambda(i, 1, p) = \lambda(i, p)$

$$\boxed{a(p, n) = \sum_{i=1}^{r-1} \lambda(i, p)\, z_i^n} \qquad \begin{array}{l} 1 \leqslant p \leqslant r-1 \\ 1 \leqslant n \leqslant r-1 \end{array} \qquad \text{(Eq8)}$$

Nous sommes donc ramenés à $(r-1)$ systèmes de $(r-1)$ équations pour déterminer les λ.

Pour p fixé nous avons $r-1$ équations à $r-1$ onconnues, qui peuvent se résoudre de façon simple :

Multiplions la n-ième équation par $\frac{1}{n!}\ \ell_i^{(n)}(0)$, où $\ell_i^{(n)}(0)$ désigne la valeur de la dérivée n-ième du polynome ℓ_i au point 0.

En ajoutant alors membre à membre les $r-1$ équations nous obtenons :

$$\sum_{n=1}^{r-1} \left(\frac{1}{n!} \ell_i^{(n)}(0) \sum_{j=1}^{r-1} \lambda(j, p)\, z_j^n \right) = \ell_i^{(p)}(0) \times \frac{1}{p!} \left(= \sum_{n=1}^{r-1} \frac{1}{n!} \ell_i^{(n)}(0)\, \delta_{pn} \right)$$

c'est-à-dire

$$\sum_{j=1}^{r-1} \lambda(j\,,\,p)\left(\sum_{n=1}^{r-1} \frac{1}{n!}\,\ell_i^{(n)}(0)\,z_j^n\right) = \ell_i^{(p)}(0) \times \frac{1}{p!}$$

De la formule de Taylor pour les polynomes de degré r − 1, on obtient :

$$\sum_{j=1}^{r-1} \lambda(j\,,\,p)\,\ell_i(z_j) = \ell_i^{(p)}(0) \times \frac{1}{p!}$$

D'où

$$\boxed{\lambda(i\,,\,p) = \ell_i^{(p)}(0) \times \frac{1}{p!}}$$

C – Forme de la matrice limite dans le cas où la matrice est scindée

L'équation 7bis s'écrit :

$$E_n = \sum_{p=1}^{r-1}\left(\sum_{i=1}^{r-1} \frac{\ell_i^{(p)}(0)}{p!} z_i^n\right) E_p$$

c'est-à-dire que M_{n+1} est alors égal à :

$$M_{n+1} = M_1 + \sum_{k=1}^{n}\left(\sum_{p=1}^{r-1}\left(\sum_{i=1}^{r-1}\left(\frac{\ell_i^{(p)}(0)}{p!} z_i^k\right)\right) E_p\right)$$

$$= M_1 + \sum_{p=1}^{r-1}\left(\sum_{i=1}^{r-1} \frac{\ell_i^{(p)}(0)}{p!} \sum_{k=1}^{n} z_i^k\right) E_p$$

Soit, à la limite, puisque $|z_i| < 1$, pour $i = 1\,,2\,,\ldots,r-1$

$$\boxed{M_\infty = M_1 + \sum_{i=1}^{r-1}\left(\sum_{i=1}^{r-1} \frac{1}{p!}\,\frac{\ell_i^{(p)}(0)}{1 - z_i}\right) E_p} \qquad \text{(Eq9)}$$

Nous allons montrer que M_∞ peut s'exprimer uniquement en fonction de coefficients B_p ou A_p .

Théorème 3 : *Dans le cas où* C *est scindée, et sous les hypothèses du théorème* 1, *on a :*

$$\boxed{M_\infty = M_1 + \frac{1}{1 + \sum_{p=1}^{r-1} B_p}\left(\sum_{k=1}^{r-1}\left(1 + \sum_{n=k+1}^{r-1} B_n\right) E_k\right)}$$

Preuve : Soit $P_i(z) = \dfrac{P(z)}{z - z_i}$

D'après la formule d'interpolation de Lagrange, tout polynome K de degré inférieur ou égal à r − 1 peut s'écrire sous la forme :

$$K(z) = \sum_{i-0}^{r-1} \frac{P_i(z)}{P_i(z_i)} K(z_i) \quad (\text{ici } z_0 = 1)$$

En prenant pour K le polynome *unité* nous obtenons :

$$1 = \sum_{i=0}^{r-1} \frac{P_i(z)}{P_i(z_i)}$$

soit, en dérivant n fois :

$$\boxed{0 = \sum_{i=0}^{r-1} \frac{P_i^{(n)}(z)}{P_i(z_i)}}$$

Remarquons que, puisque

$$P(z) = \sum_{i=0}^{r} (z - z_i) = \left(\sum_{i=1}^{r-1} (z - z_i)\right)(z - 1)$$

et que

$$\ell_i(z) = \prod_{j \neq i} \frac{z - z_j}{z_i - z_j}$$

pour $1 \leqslant i \leqslant r - 1$

$$(1 - z)\frac{\ell_i(z)}{1 - z_i} = \frac{P_i(z)}{P_i(z_i)}$$

$$\frac{P_i'(z)}{P_i(z_i)} = -\frac{\ell_i(z)}{1 - z_i} + (1 - z)\frac{\ell_i'(z)}{1 - z_i}$$

De façon générale, on aurait :

$$(1 - z)\frac{\ell_i^{(n)}(z)}{1 - z_i} = \frac{P_i^{(n)}(z)}{P_i(z_i)} + n\frac{\ell_i^{(n-1)}(z)}{1 - z_i}$$

Donc

$$\frac{\ell_i^{(n)}(0)}{1 - z_i} = \frac{P_i^{(n)}(0)}{P_i(z_i)} + n\frac{P_i^{(n-1)}}{P_i(z_i)} \ldots + n!\frac{P_i(0)}{P_i(z_i)}$$

Alors

$$\sum_{i=1}^{r-1} \frac{\ell_i^{(k)}(0)}{(1 - z_i)\,k\,!} = \sum_{i=0}^{r-1} \frac{1}{k\,!}\left(\frac{P_i^{(k)}(0)}{P_i(z_i)} \ldots + k\,!\,\frac{P_i(0)}{P_i(z_i)}\right)$$

$$- \frac{1}{k\,!}\left(\frac{P_0^{k}(0)}{P_0(1)} \ldots + k!\,\frac{P_0(0)}{P_0(1)}\right)$$

Or $P_0^{(n)}(0) = n!\ B_n$

Donc

$$\sum_{i=1}^{r-1} \frac{\ell_i^{(k)}(0)}{(1 - z_i)\, k!} = 1 - \frac{\sum_{n=0}^{k} B_n}{1 + \sum_{n=0}^{r-1}} = \frac{1 + \sum_{n=k+1}^{r-1} B_n}{1 + \sum_{n=0}^{r-1} B_n}$$

Ce qui nous donne :

$$\boxed{M_\infty = M_0 + \frac{1}{1 + \sum_{n=0}^{r-1} B_n} \left(\sum_{k=1}^{r-1} \left(1 + \sum_{n=k+1}^{r-1} B_n\right) E_k \right)}$$

V – CAS DES RACINES D'INDICE QUELCONQUE

A – Préliminaires

Nous sommes toujours dans les hypothèses du théorème 1. Soit

$$z_0 = 1 \quad \text{et} \quad z_1, z_2, \ldots, z_p$$

les autres racines de l'équation minimale, respectivement d'indice $\beta_1\ \beta_2, \ldots \beta_p$

avec $\sum_{i=1}^{p} \beta_i = r - 1$

Soit $P(X)$ le polynome minimal de C

Nous noterons

$$R(X) = \prod_{i=1}^{p} (X - z_i)^{\beta_i} = (X^r + B_{r-1} X^{r-1} \ldots + B_0)/(X - 1)$$

$$\ell_i(z) = \prod_{j \neq i} (z - z_j)^{\beta_j} \frac{1}{\prod_{j \neq i} (z_i - z_j)^{\beta_j}}$$

– $P(X) = (1 - X)\, R(X)$ et à partir de R on définira de même

$$\left[\begin{array}{l} i \neq 0 \quad L_i(z) = \ell_i(z) \times \dfrac{1 - z}{1 - z_i} \\[2ex] (i = 0) \quad L_0(z) = \dfrac{R(z)}{R(1)} \end{array}\right.$$

Rappel : La base de Lagrange pour l'interpolation polynomiale d'Hermite est l'ensemble des polynomes

$$(f_{i,j}(z)\ 1 \leqslant i \leqslant p \quad \text{et} \quad 0 \leqslant j \leqslant \beta_1 - 1$$

vérifiant les conditions

$$\forall j \ \ 0 \leqslant j \leqslant \beta_{i-1} \ \ \forall k \neq i \ \ \forall 0 \leqslant m \leqslant \beta_k - 1 \left[\begin{array}{l} f_{i,j}^{(m)}(z_k) = 0 \\ f_{i,j}^{(m)}(z_i) = \delta_{jm} \end{array}\right.$$

Tout polynome de degré inférieur ou égal à r – 1 s'écrit alors :

$$T(z) = \sum_{i=1}^{p} \left(\sum_{j=0}^{\beta_i - 1} f_{ij}(z)\, T^{(j)}(z_i) \right)$$

Un calcul assez simple, donné en annexe, montre que :

$$f_{i,j}(z) = \left| \sum_{n=j}^{i-1} C_h^j \left(\frac{1}{\ell_i}\right)^{(n-j)}(z_i) \quad (z - z_i)^n \right| \ell_i(z)$$

Si a(n , q) est le coefficient de E_q dans l'expression de E_n nous avons, en prenant comme base de l'espace vectoriel des suites solutions de l'équation de récurrence

$$u_{i,j}(n) = \frac{n!}{(n-j)\,!}\, z_i^n \quad 0 \leqslant i \leqslant p \qquad \text{si} \quad n < j \ \frac{n!}{n-j\,!} = 0$$

$$0 \leqslant j \leqslant \beta_i - 1$$

$$a_{(n,q)} = \sum_{i=1}^{p} \left(\sum_{j=0}^{\beta_i - 1} (\chi_{i,j,q}) \frac{n!}{(n-j)!}\, z_i^n \right)$$

Les λ(i , j , q) étant déterminés par le système de r équations à r inconnues

$$0 \leqslant k \leqslant N - 1 \left(\sum_{i=1}^{p} \left(\sum_{j=0}^{\beta_i - 1} \lambda(i, j, q) \frac{k\,!}{(k-j)\,!}\, z_i^k \right) = \delta_{kq} \right.$$

Multiplions la $k^{\text{ème}}$ équation par $\frac{1}{k!} f_{i,j}^{(k)}(0)$, et additionnons les équations. D'après les conditions imposées aux $f_{i,j}$, et la base choisie, le coefficient de λ(m , n , q) est $z_m^n \times f_{i,j}^{(n)}(z_m)$

Donc seul le terme en λ(i , j , q) a un coefficient non nul, égal à z_i^j .

D'où :

$$\lambda(i, j, q) = \frac{1}{q!} \frac{f_{i,j}^{(q)}(0)}{z_i^j}$$

B – Détermination de la matrice limite des M_n

En remarquant que pour $|z_i| < 1$ (ce qui est supposé vrai ici).

$$\sum_{n=0}^{\infty} \frac{n!}{(n-j)!} z_i^n \, d_{z_i}^j \left(\frac{1}{1-z}\right)^{(16)}$$

Nous obtenons :

$$M_\infty = M_o + \sum_{q=0}^{\beta-1} \frac{1}{q!}\left(\sum_{i=1}^{p} \sum_{j=0}^{\beta_i - 1} d_{z_i}^j\right)\left(\frac{1}{1-z}\right) f_{i,j}^{(q)}(0)\Big) E_q$$

En considérant l'interpolation d'Hermite par rapport au point z_i d'ordre β_i et $z_0 = 1$ l'ordre 1, nous avons une formule semblable à celle donnée dans le rappel, où nous noterons les polynomes d'interpolation avec des lettres majuscules.

En particulier pour le polynome constant égal à 1 nous avons :

$$1 = \sum_{i=1}^{p} F_{i,0}(z) + F_{0,0}(z).$$

avec

$$F_{i,0}(z) = \sum_{p=0}^{\beta_1 - 1} \left(\frac{1}{L_i}\right)^{(p)}_{(z_i)} (z - z_i)^p \; L_i(z)$$

et

$$F_{0,0}(z) = R(z)$$

Pour trouver M_∞ nous allons tout d'abord démontrer que :

$$\boxed{(1-z) \sum_{j=0}^{\beta_i - 1} \left(\frac{1}{1-z}\right)^{(j)}_{(z_i)} f_{i,j}(z) = F_{i,0}(z)}^{(17)}$$

En effet :

$$\ell_i(z)\,(1-z) \sum_{j=0}^{\beta_i - 1} \left(\frac{1}{1-z}\right)^{(j)}(z_i) \sum_{m=j}^{\beta_i - 1} C_m^j \left(\frac{1}{i}\right)^{(m-j)}_{(z_i)} (z - z_i)^m$$

$$= (1-z)\,\ell_i(z) \sum_{m=0}^{\beta_i - 1} \left[\sum_{j=0}^{m} C_m^j \left(\frac{1}{\ell_i}\right)^{(m-j)}_{(z_i)} \left(\frac{1}{1-z}\right)^{(j)}_{(z_i)} (z - z_i)^m \right] \quad (17)$$

---

(16) $d_{z_i}^j \left(\frac{1}{1-z}\right)$ représente la dérivée d'ordre j de la fonction $1 \to \frac{1}{1-z}$ au point z_i .

(17) $\left(\frac{1}{1-z}\right)^{(j)}_{(z_i)}$ ou $d_{z_i}^j \left(\frac{1}{1-z}\right)$

Or d'après la formule de Leibnitz

$$\sum_{j=0}^{m} C_m^j \left(\frac{1}{\ell_i}\right)^{(m-j)}_{(z_i)} \times \left(\frac{1}{1-z}\right)^{(j)}_{(z_i)} = \left(\frac{1}{\ell_i \times (1-z)}\right)^{(m)}_{(z_i)} = \left(\frac{1}{L_i}\right)^{(m)}_{(z_i)} \times \frac{1}{1-z_i}$$

D'où

$$(1-z) \sum_{j=0}^{\beta_i - 1} \left(\frac{1}{1-z}\right)^{(j)} (z_i)\, f_{i,j}(z) = \frac{1-z}{1-z_i}\, \ell_i(z) \sum_{m=0}^{\beta_i - 1} \left(\frac{1}{L_i}\right)^{(m)} (z_i) \times (z - z_i)^m$$

c.q.f.d

Un simple raisonnement par récurrence montre que :

$$\sum_{j=0}^{\beta_i = 1} \left(\frac{1}{1-z}\right)^{(j)} (z_i)\, f_{ij}^{(k)}(0) = F_{i,0}^{(k)} + \ldots + k!\ F_{i,0}(0) = \sum_{\ell=0}^{k} \frac{k!}{\ell!} F_{i,0}^{(\ell)}(0)$$

D'où

$$\frac{1}{q!} \sum_{i=1}^{p} \sum_{j=0}^{\beta_i - 1} \left(\frac{1}{1-z}\right)^{(j)} (z_i)\, f_{ij}^{(q)}(0) = \sum_{i=1}^{p} \sum_{\ell=0}^{q} \frac{1}{\ell!} F_{i,0}^{(\ell)}(0)$$

$$= \sum_{\ell=0}^{q} \frac{1}{\ell!} \sum_{i=1}^{p} F_{i,0}^{(\ell)}(0)$$

Or :

$$\sum_{i=1}^{p} F_{i,0}^{(\ell)}(0) = \delta_{1\ell} - F_{0,0}^{(\ell)}(0) \quad \text{avec} \quad F_{0,0}^{(\ell)}(0) = \ell!\ B_\ell \times \frac{1}{R(1)}$$

D'où :

$$\frac{1}{q!} \sum_{i=1}^{p} \sum_{j=0}^{\beta_i - 1} \left(\frac{1}{1-z}\right)^{j} f_{ij}^{(q)}(0) = 1 - \sum_{\ell=0}^{q} \frac{B\ell}{R_{(1)}} = \frac{1}{R_1} + \sum_{\ell=q+1}^{r-1} \frac{B\ell}{R_{(1)}}$$

ce qui donne, en posant $B_r = 1$

$$= \frac{\sum_{\ell=q+1}^{r} B_e}{\sum_{\ell=0} B_e}$$

On retrouve donc, comme dans le cas des racines distinctes

$$M_\infty = M_0 + \sum_{q=0}^{r-1} \frac{\sum_{\ell=q+1}^{r} B_e}{\sum_{\ell=0}^{r} B_e} E_q$$

Soit, en revenant à la définition des E_q

$$M_\infty = \frac{1}{\sum_{\ell=0}^{r} B_e} (C^r + B_{r-1} C^{r-1} \ldots + B_0 I) M_0$$

$$\boxed{M_\infty = \frac{1}{\sum_{\ell=0}^{r} B_e} R(C) \, . \, M_0}$$

(ce qui vérifie bien $(C - I) M_\infty = 0$
soit $CM_\infty = M_\infty$ comme on s'en doutait)

C – Annexe

Détermination des $f_{i,j}$ (z). Le premier système de conditions entraine que f_{ij} (z) est divisible par $\ell_i(z)$.

Soit $\quad f_{i,j}(z) = A_{i,j}(z) \ell_i(z)$

où $\quad A_{i,j}$ est un polynome de degré α_{i-1}

Pour déterminer $A_{i,j}(z)$, il suffit d'utiliser la deuxième série de conditions et la formule de Newton. En effet :

$$A_{i,j}(z) = \frac{f_{i,j}(z)}{\ell_i(z)}$$

Soit :

$$A_{i,j}^{(p)}(z_i) = 0 \quad \text{si} \quad p < j$$

$$A_{i,j}^{(p)}(z_i) = C_p^j \left(\frac{1}{\ell_i(z)}\right)^{(p-j)} (z_i) \qquad \text{(Formule de Leibnitz)}$$

Donc

$$A_{i,j}(z) = \left(\sum_{p=j}^{\alpha_i - j} C_p^j \left(\frac{1}{\ell_i(z)}\right)^{(p-j)} (z_i) \times (z - z_i)^p\right)$$

Remarques

1/ **Théorème**[18] : *Si* 1 *est valeur propre simple la matrice limite* M_∞ *quand elle existe, à toutes ses lignes identiques.*

Démonstration : Puisque $M_\infty = CM_\infty$, toute colonne de M_∞ peut être considérée comme vecteur propre[19] de C associé à la valeur propre 1. Dans le cas ou 1 est valeur propre[19] simple, les vecteurs propres sont de la forme $\begin{bmatrix} x \\ x \\ \vdots \\ \cdot \end{bmatrix}$, d'où le résultat.

---

(18) Ce théorème, extrêmement simple, nous apparaît *fondamental* puisque cela signifie que, dans le cas (très fréquent) où 1 est valeur propre simple M_n et $M_1 \, . \, M_2 \ldots M_n$ ont la même limite lorsque $n \to \infty$. Ce résultat n'avait pu être montré par Harary-Lipstein-Styan.

En particulier dans le cas de deux états si $C \neq I$ et si M_n tend vers une limite stochastique, M_∞ peut être associée à un processus homogène *stationnaire* (qu'on pourrait appeler *"tangent"* au processus non homogène initial).

Remarque 2

Si M *est régulière les* $(E_i)_{i=1,\ldots r-1}$ *sont indépendantes.*

Remarque 3

Cas où les $(E_i)_{i=1\ldots r-1}$ sont dépendantes (cf. 2 exemples ci-après)

Alors, on peut écrire, les E_i étant indépendantes,

$$i = 1, \ldots, p-1 \quad \text{avec} \quad p < r : (C^p + a_{p-1} C^{p-1} \ldots + a_1 I) M = 0$$

le polynome $C^p + a_{p-1} C^{p-1} \ldots + a_1 I$ *divise le polynome minimal* (1) et pourra être appelé *polynome causatif* de M ; il faut et il suffit alors de raisonner sur le polynome causatif de M (1 est toujours racine du polynome causatif).

Démonstration de (1) : soit $\lambda_1, \lambda_2 \ldots \lambda_{p-1}, \mu$ les racines du polynome causatif de M. Si par exemple μ n'était pas valeur propre de C alors $\exists$ un vecteur colonne de M, note M_α, tel que :

$$\prod_{i-1}^{p-1} (C - \lambda_i I) M_\alpha \neq 0 \quad \text{et} \quad (C - \mu I) \pi (C - \lambda_i I) M_\alpha = 0$$

$$= M_\beta \qquad \Rightarrow (C - \mu I) M_\beta = 0$$

$$\Rightarrow \mu \text{ valeur propre de } C$$

c.q.f.d.

Exemple 1 : Si $C = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ -1 & 0 & 2 \end{bmatrix}$ Prenons $M = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a & b & c \end{bmatrix}$

alors le polynome causatif est $(C - I)(C - \frac{1}{2} I) . M = 0$

i.e : $$C^2 M - CM = \frac{1}{2}(CM - M) \Rightarrow E_2 = \frac{1}{2} E_1$$

$C^n M$ converge, bien que C^n ne converge pas.

Dans le cas où les $(E_i)_{i=1\ldots r-1}$ sont dépendantes le polynôme causatif joue le rôle de polynôme minimal. On a donc l'extension du Théorème 1.

Théorème 1bis : *La suite* $(M_n)_{n \geq 1}$ *est convergente ssi les conditions suivantes sont satisfaites :*

- - - - - - - - - - - - -

(19) à droite

c_1 : 1 *est d'indice* 1 *dans le polynôme causatif*

c_2 : *les autres racines du polynome causatif sont de module strictement inférieur à* 1 .

Remarque 4 (Exemple 2)

Dans le cas de deux états le cas particulier se réduit au cas où M_1 est non régulière c'est-à-dire M_1 stationnaire.

Remarque 5

Si 1 est valeur propre à spectre simple et si toutes les valeurs propres de C apparaissent dans le "polynome causatif" de M_1 la limite M_∞ conserve la même valeur que celle trouvée dans le cas où le "polynome causatif" de M_1 était le polynome minimal de C.

Les conditions sont nécessaires pour que le passage à la limite fait pour le polynome minimal soit licite (on peut évidemment étendre ceci au cas où 1 est valeur propre à spectre simple et où toutes les valeurs propres de C sont de module strictement inférieur à 1).

En effet on a $M_\infty = \dfrac{K(c)}{K(1)} M_1$. où K est le "polynome causatif".

Montrons que R étant le polynome minimal de C./C – I

$\left(\dfrac{R(C)}{R(1)} - \dfrac{K(C)}{K(1)}\right) M_0 = 0$ On peut écrire $R(C) = K(C) \times B(C)$.

Ce qui revient donc à démontrer que :

$$\frac{K(C)}{K(1)}\left(B(C)\frac{K(1)}{R(1)} - I\right) M_0 = 0$$

c'est-à-dire que C – I est en facteur dans le crochet.

or $$B(1)\frac{K(1)}{R(1)} = \frac{R(1)}{K(1)} \times \frac{K(1)}{R(1)} = 1 \qquad \text{c.q.f.d.}$$

Donc, si le polynome causatif fait intervenir toutes les valeurs propres de C, il suffit d'utiliser les résultats obtenus dans le cas général. On peut même par un raisonnement analogue à celui fait dans la remarque 2 utiliser le polynome caractéristique de C divisé par C – I à la puissance β_0.

Remarque 6

Pour que $1, \lambda_1, \lambda_2 \ldots \lambda_p$ soient les valeurs propres de C qui interviennent dans le polynome causatif il faut et il suffit que si $F_0, F_1, \ldots F_{p-1}$ sont les sous espaces propres associés à ces valeurs propres tous les vecteurs colonnes de M soient des vecteurs de $F_0 \oplus F_1 \oplus \ldots \oplus F_p$.

2ème PARTIE

DETERMINATION ALGORITHMIQUE[20] DES MATRICES DE CAUSALITE DANS LE CAS DES CHAINES A DEUX ETATS

I – RAPPELS :

Nous avons, avec les notations de la 1ère partie, en remplaçant E_1 par E :

$$(S)\quad \begin{cases} P_1 = P_0 M_1 \\ P_2 = P_1 M_2 = P_1 (M_1 + E) \\ \cdot \\ \cdot \\ \cdot \\ P_{n+1} = P_n M_{n+1} = P_n \; M_n + \dfrac{E}{k^{n-1}} \end{cases}$$

(k étant la 2ème valeur propre de la matrice de causalité C)

La matrice E est dite "d'écart d'homogénéité", avec :

$$E = \begin{pmatrix} \epsilon_1 & -\epsilon_1 \\ \epsilon_2 & -\epsilon_2 \end{pmatrix} = M_1 (C - I) \qquad \left(C = \begin{pmatrix} c_1 & 1 - c_1 \\ c_2 & 1 - c_2 \end{pmatrix} \right)$$

Nous avons démontré que nous devions avoir, pour avoir *convergence,* donc pour que le modèle puisse refléter une situation *réelle, soit :*

i) $|c_1 - c_2| < 1$

ii) $M_{n+1} = M_n + \dfrac{E}{k^{n-1}}$, avec $k = \dfrac{1}{c_1 - c_2} = 1 + \dfrac{\epsilon_2 - \epsilon_1}{m_2 - p_2} = \dfrac{m_1 - p_1}{m_2 - p_2}$

soit : (iii) $c_1 - c_2 = \pm 1$, avec certaines conditions restrictives.

Nous allons montrer que, *en général, la seule connaissance des parts de marché des 6 premières périodes suffit pour déterminer les matrices de transition.* Les cas particuliers seront étudiés dans le paragraphe II de cette 2ème partie, et seront essentiellement des *points techniques, non essentiels.*

II – SYSTEME FONDAMENTAL D'EQUATIONS : EXISTENCE D'UNE SOLUTION

Les 4 premières équations, par exemple, du système S peuvent s'écrire :

---

(20) Nous avons ici une détermination analytique (et non statistique) c'est-à-dire que nos modèles sont supposés parfaits.

$$(S_1)\quad \begin{cases} \text{(I)} & \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = \begin{pmatrix} P_0 \\ P_1 \end{pmatrix} M_1 + \begin{pmatrix} 0 \\ P_1 \end{pmatrix} E, \quad \text{avec} \quad 0 = (0\,,0) \\ \text{(II)} & \begin{pmatrix} P_2 \\ P_3 \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} (M_1 + E) + \begin{pmatrix} 0 \\ P_2 \end{pmatrix} E/k \\ \text{(III)} & \begin{pmatrix} P_3 \\ P_4 \end{pmatrix} = \begin{pmatrix} P_2 \\ P_3 \end{pmatrix} \left(M_1 + E + \frac{E}{k}\right) + \begin{pmatrix} 0 \\ P_3 \end{pmatrix} E/k^2 \end{cases}$$

A – Cas où $\boxed{P_0 \neq P_1 \neq P_2 \neq P_3 \neq P_4 \neq P_5}$

On peut alors écrire, puisque $\begin{pmatrix} P_0 \\ P_1 \end{pmatrix}$ est *inversible :*

$$M_1 = \begin{pmatrix} P_0 \\ P_1 \end{pmatrix}^{-1} \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} - \begin{pmatrix} P_0 \\ P_1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ P_1 \end{pmatrix} E$$

ou, en posant

$$\boxed{T_1 = \begin{pmatrix} P_0 \\ P_1 \end{pmatrix}^{-1} \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}; \quad F_1 = \begin{pmatrix} P_0 \\ P_1 \end{pmatrix}^{-1}},$$

et, plus généralement, lorsque de telles matrices existent :

$$T_n = \begin{pmatrix} P_{n-1} \\ P_n \end{pmatrix}^{-1} \begin{pmatrix} P_n \\ P_{n+1} \end{pmatrix}; \quad F_n = \begin{pmatrix} P_{n-1} \\ P_n \end{pmatrix}^{-1}$$

$$M_1 = T_1 - F_1 \begin{pmatrix} 0 \\ P_1 \end{pmatrix} E$$

$$M_1 + E = T_2 - F_2 \begin{pmatrix} 0 \\ P_2 \end{pmatrix} E/k$$

$$M_1 + E + \frac{E}{k} = T_3 - F_3 \begin{pmatrix} 0 \\ P_3 \end{pmatrix} E/k^2$$

$$M_1 + E\left(\frac{1}{k} + \frac{1}{2}\right) = T_4 - F_4 \begin{pmatrix} 0 \\ P_4 \end{pmatrix} E/k^3$$

On peut alors montrer, ce que nous ferons dans le paragraphe III, que ce système détermine de façon unique M_1 et E.

Nous allons maintenant envisager tous les cas particuliers :

B) Cas où $P_0 = P_1$

On a donc $P_0 = P_0 M_1$, c'est-à-dire que P_0 est *vecteur propre* à gauche de M_1 pour la valeur propre 1, c'est-à-dire que P_0 est *stationnaire* pour le processus qui *serait homogène* de matrice constante M_1 : cela *entraîne*(21), en particulier, que l'on ait, en notant $P_0 = (a_0\,, 1 - a_0)$, M_1 de la forme :

---

(21) Ce résultat, *assez important,* sera démontré par la suite.

$$\lambda_1 \begin{pmatrix} a_0 & 1-a_0 \\ a_0 & 1-a_0 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{avec} \quad \lambda_1 + \mu_1 = 1 \quad \text{et} \quad \lambda_1 \mu \geqslant 0$$

B_1) Cas où $P_0 = P_1 = P_2$

On a donc, de même, M_2 de la forme $\lambda_2 \begin{pmatrix} a_0 & 1-a_0 \\ a_0 & 1-a_0 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

avec $\lambda_2 + \mu_2 = 1$ et $\lambda_2, \mu_2 \geqslant 0$

On a ainsi, (toujours avec $\lambda_2 + \mu_2 = \lambda_1 + \mu_1 = 1$)

$$m_2 = \lambda_2 a_0 + \mu_2$$

$$P_2 = \lambda_2 a_0$$

$$m_1 = \lambda_1 a_0 + \mu_1$$

$$P_1 = \lambda_1 a_0$$

$$k = \frac{\mu_2}{\mu_1}, \ \epsilon_1 (\lambda_2 - \lambda_1)(a_0 - 1), \ \epsilon_2 = (\lambda_2 - \lambda_1) a_0$$

On va alors démontrer le résultat suivant :

Lemme : *Si* M_1 *et* $M_2 = C.M_1$ *ont même vecteur propre à gauche (pour la valeur propre* 1*) alors il en est de même pour* $M_3 = C^2 M_1$

Preuve : Il suffit de montrer que M_3 est de la forme

$$\lambda_3 \begin{pmatrix} a_0 - 1 - a_0 \\ a_0 - 1 - a_0 \end{pmatrix} + \mu_3 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{avec} \quad \lambda_3 + \mu_3 = 1 \quad \text{et} \quad \lambda_3, \mu_3 \geqslant 0$$

Pour cela calculons m_3 et p_3. Par hypothèse, nous avons :

$$m_3 = m_2 + \frac{\epsilon_1}{k} = (\lambda_2 a_0 + \mu_2) + \frac{\mu_1}{\mu_2}((\lambda_2 - \lambda_1)(a_0 - 1))$$

$$= a_0 (\lambda_2 + \frac{\mu_1}{\mu_2}(\lambda_2 - \lambda_1)) + \mu_2 - \frac{\mu_1}{\mu_2}(\lambda_2 - \lambda_1)$$

en posant

$$\lambda_3 = \lambda_2 + \frac{\mu_1}{\mu_2}(\lambda_2 - \lambda_1) \quad \text{et} \quad \mu_3 = \mu_2 - \frac{\mu_2}{\mu_2}(\lambda_2 - \lambda_1)$$

on a bien :

$$\boxed{\lambda_3 + \mu_3 = \lambda_2 + \mu_2 \ (= 1)}$$

et on peut montrer aussi soit par un raisonnement direct (M_3 devant être stochastique, μ_3 doit être nécessairement $\leqslant 1$, donc λ_3 doit être $\geqslant 0$), soit en développant les expressions obtenues, que

$$\boxed{\lambda_3 \ , \mu_3 \geqslant 0}$$

c.q.f.d.

Le lemme *entraîne en particulier* puisque P_2 est vecteur propre pour M_3, que $\boxed{P_3 = P_2}$. Le même lemme montrerait *évidemment* que $\boxed{P_{n+1} = P_n}$, pour tout n.

Remarque : Bien entendu, le cas $P_0 = P_1 = P_2$ (peu fréquent dans la pratique !) peut s'étudier directement plus simplement (cependant, le lemme étant intéressant en lui-même nous avons préféré donner l'approche précédente).

En effet :

$$P_2 = P_1 \Rightarrow \boxed{P_0 E = 0}$$

$$\begin{aligned} P_3 = \underset{\substack{\uparrow \\ \text{(hypoth.)}}}{P_2} M_3 = P_0 M_3 = P_0\left(M_1 + E + \frac{E}{k}\right) &= P_0 + P_0 E + P_0 \frac{E}{k} \\ &= P_0 + 0 + 0 \\ &= P_0 \end{aligned}$$

c.q.f.d.

B_2 $\boxed{P_2 \neq P_1}$ (avec $P_2 = P_1 (M_1 + E) = P_0 (M_1 + E) = P_0 + P_0 E$)

On a, en particulier, $\boxed{P_0 E \neq 0}$

On peut alors écrire :

$$\begin{aligned} P_3 = P_2 M_3 &= (P_0 + P_0 E)\left(M_1 + E + \frac{E}{k}\right) \\ &= P_0 + P_0 E + P_0 E\left(M_1 + \frac{I}{k} + E\left(1 + \frac{1}{k}\right)\right) \\ &= P_2 + P_0 E\left(M_3 + \frac{I}{k}\right) \end{aligned}$$

On a le résultat suivant :

Dans le cas $\boxed{P_0 = P_1 \neq P_2 = P_3}$, on a 2 possibilités :

où la suite des vecteurs d'états est de la forme

$$P_0 \ , P_0 \ , P_2 \ , P_2 \ , P_0 \ , P_0 \ , P_2 \ , P_2 \ , P_0 \ldots$$

où cette suite devient stationnaire à partir de la période 2.

Ce résultat, *technique,* est démontré ci-dessous.

Deux cas sont possibles :

$$\boxed{B_2 \text{ ⓐ} : P_3 = P_2} \Rightarrow P_0 E \; M_3 + \frac{I}{k} = 0$$

$$\Rightarrow \left[\begin{array}{l} P_0 E \text{ vecteur propre (puisque } P_0 E \neq 0) \\ \text{de } M_3 + \frac{I}{k} \text{ pour la valeur propre } 0 \end{array}\right.$$

$$\Rightarrow M_3 + \frac{I}{k} \text{ n'est pas inversible}$$

$$\Longleftrightarrow \left[\begin{array}{l} \text{Il existe } \gamma \left(\text{ou } \frac{1}{\gamma}\right) \neq 0 \text{ tel que :} \\ \quad \left[\begin{array}{l} P_3 = \gamma \left(m_3 + \frac{1}{k}\right) \\ 1 - p_3 + \frac{1}{k} = \gamma (1 - m_3) \end{array}\right. \end{array}\right.$$

$$\Rightarrow 1 - \gamma m_3 - \frac{\gamma}{k} + \frac{1}{k} = \gamma - \gamma m_3$$

$$\Rightarrow (1 - \gamma) \left(1 + \frac{1}{k}\right) = 0$$

$$\Longleftrightarrow \boxed{\gamma = 1} \quad \text{ou} \quad \boxed{k = -1}$$

Etudions le cas : $k = -1$

On a alors :

$$M_2 = M_1 + E$$

$$M_3 = M_1 + E - E = M_1$$

L'équation $P_3 = P_2 M_3$ s'écrit : $P_3 = P_2 M_1$ et on a donc, par hypothèse : $P_2 = P_3 = P_2 M_1$, c'est-à-dire que P_2 est *vecteur propre,* à gauche, de M_1. Comme on avait aussi P_0 *vecteur propre,* à gauche, de M_1, la condition $P_0 \neq P_2$ entraîne que M_1 a $\lambda = 1$ comme valeur propre double : M_1 étant stochastique cela signifie que : $\boxed{M_1 = I}$

Puisque $C = M_2 \cdot M_1^{-1}$, on a $C = M_2$, et C est donc *stochastique :* comme "-1" est valeur propre de C (en effet $k = -1$), cela que :

$$\boxed{C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}$$

Etudions le cas : $\gamma = 1$

$$\Rightarrow \boxed{m_3 - p_3 = -\frac{1}{k}}$$

Si la chaîne *n'est pas homogène*[22] on a aussi :

$$\frac{m_3 - p_3}{m_2 - p_2} = \frac{1}{k} = \frac{m_2 - p_2}{m_1 - p_1} \qquad \text{(voir 1ère partie)}$$

$$\Rightarrow m_2 - p_2 = -1$$

Comme M_2 est stochastique, cela signifie que $\boxed{m_2 = 0}$ et $\boxed{p_2 = 1}$

$M_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ entraîne que $(P_3) = P_2 = (a_2 \ , \ 1 - a_2)$

$$= (a_1 \ , \ 1 - a_1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$= (1 - a_1 \ , \ a_1)$$

$$= \mathit{1} - P_1 = \mathit{1} - P_0$$

(avec $\mathit{1} = (1 \ , \ 1)$)

On a évidemment aussi :

$$\boxed{\begin{aligned} \epsilon_1 &= -m_1 \\ \epsilon_2 &= 1 - p_1 \end{aligned}} \qquad (1)$$

La relation $M_3 = C\,M_2$ entraîne que l'on a, puisque $M_2^{-1} = M_2$,

$$\boxed{C = \begin{pmatrix} 1 - m_3 & m_3 \\ 1 - p_3 & p_3 \end{pmatrix}} \qquad (2)$$

On a enfin $c_1 = 1 + \dfrac{\epsilon_3}{m_2 - p_2}$ et $c_2 = \dfrac{\epsilon_4}{m_2 - p_2}$ qui donnent, avec $\epsilon_3 = \dfrac{\epsilon_1}{k}$ et $\epsilon_4 = \dfrac{\epsilon_2}{k}$, c'est-à-dire :

$$\left[\begin{aligned} &\epsilon_3 = m_1 (m_3 - p_3) \\ &\text{et} \\ &\epsilon_4 = (p_1 - 1)(m_3 - p_3) \end{aligned}\right.$$

$$\boxed{\begin{aligned} c_1 &= 1 + \left(\frac{m_3 - p_3}{m_2 - p_2}\right) \times (m_1) \\ c_2 &= (p_1 - 1)\frac{m_3 - p_3}{m_2 - p_2} \end{aligned}}$$

---

(22) *Remarque :* Si la chaîne est homogène on a $M_3 = M_2 = M_1$ donc $P_0 = P_1$ devrait entraîner $P_1 = P_2$: comme c'est faux, par hypothèse, *la chaîne ne peut être homogène.*

Des calculs simples, mais fastidieux, montrent alors que l'on a, nécessairement :

$$m_1 = m_3 = 1$$

$$p_1 = p_3 = 0$$

ce qui, avec $m_2 = 0$ et $p_2 = 1$, donne :

$$M_1 = M_3 = M_5 = \ldots = M_{2n+1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$M_2 = M_4 = \ldots = M_2 n = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (= C)$$

c.q.f.d.

B_2 (β) : $P_3 \neq P_2$ $(\neq P_1 = P_0)$[23]

On pourrait voir simplement que l'on est dans *un cas analogue* au cas A, en *décalant tout d'une période.*

III – QUELQUES EXEMPLES ILLUSTRATIFS

Cas "A" $(P_0 \neq P_1 \neq P_2 \neq P_3 \neq P_4)$

1/ $P_0 = (1\,,0)$ $\quad M_1 = \begin{pmatrix} .8 & .2 \\ .4 & .6 \end{pmatrix}$, $M_2 = \begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix}$ $\Rightarrow$ $\epsilon_1 = -.1$, $\epsilon_1 = +.1$ $\quad k = 2$

Alors $P_1 = (.8,.2)$, $P_2 = (.66,.34)$, $P_3 = (.616,.384)$, $P_4 = (.6058,.3942)$

. . .

On pourrait voir que $P_n \xrightarrow[r \to \infty]{} (.6,.4)$ et que $M_n \xrightarrow[n \to \infty]{} \begin{pmatrix} .6 & .4 \\ .6 & .4 \end{pmatrix}$

(Remarque : Le vecteur $P'_0 = (2/3\,,1/3)$ aurait été *stationnaire* pour le processus homogène de matrice, constante, M_1)

2/ $P_0 = (.7,.3)$ avec les mêmes matrices M_1 et M_2

Alors $P_1 = (.68,.32)$, $P_2 = (.636,\ .364)$ . . . , mais on aura aussi, évidemment :

$$P_n \xrightarrow[n \to \infty]{} (.6,\ .4)$$

3/ (Exemple cité par Lipstein, sans que les matrices limites soient étudiées)

---

(23) ce sera d'ailleurs, pratiquement, le cas le plus fréquent.

$$\left.\begin{array}{l} M_1 = \begin{pmatrix} .77 & .23 \\ .40 & .60 \end{pmatrix} \\ M_2 = \begin{pmatrix} .70 & .30 \\ .40 & .60 \end{pmatrix} \end{array}\right] \Rightarrow \left[\begin{array}{l} \epsilon_1 = -.07 \\ \epsilon_2 = 0 \\ k = \dfrac{.37}{.30} = \dfrac{1}{.810810810\ldots} \end{array}\right.$$

On aura alors, quelque soit P_0 , $P_n \xrightarrow[n \to \infty]{} (.4,.6)$

Si nous prenons, par exemple, $P_0 = (.4,.6)$, nous aurons

$$P_1 = (.548,.452)\ , P_2 = (.5644,.4356), \ldots$$

Cas B_1 $P_0 = P_1 = P_2$

(Exemple d'une suite stationnaire de vecteurs d'états, la suite de matrice de transitions associée étant à causalité constante, mais non homogène)[24].

$$P_0 = (.6,.4) \qquad \left.\begin{array}{l} M_1 = \begin{pmatrix} .84 & .16 \\ .24 & .76 \end{pmatrix} \\ M_2 = \begin{pmatrix} .8 & .2 \\ .3 & .7 \end{pmatrix} \end{array}\right] \Rightarrow \left[\begin{array}{l} \epsilon_1 = -.04 \\ \epsilon_2 = +.06 \\ k = 1.2 \end{array}\right.$$

On a alors $P_1 = P_2 = P_3 = \ldots = (.6,.4)$ et, *évidemment*, $M_\infty = \begin{pmatrix} .6 & .4 \\ .6 & .4 \end{pmatrix}$

Cas B_2 , β : $(P_0 = P_1 \neq P_2 \neq P_3)$

$$P_0 = (.6,.4) \qquad \left|\begin{array}{l} M_1 = \begin{pmatrix} .84 & .16 \\ .24 & .76 \end{pmatrix} \\ M_2 = \begin{pmatrix} .78 & .22 \\ .28 & .72 \end{pmatrix} \end{array}\right] \Rightarrow \begin{array}{l} \epsilon_1 = -.06 \\ \epsilon_2 = +.04 \\ k = 1.2 \end{array}$$

$$P_1 = (.6,.4)\ , P_2 = (.58,.42)\ , P_3 = (.555,\ .445)\ , \ldots$$

Et on pourrait montrer que $P_n \xrightarrow[n \to \infty]{} (.48,.52)$ et que, aussi $M_n \xrightarrow[n \to \infty]{} \begin{pmatrix} .48,.52 \\ .48,.52 \end{pmatrix}$

Remarque : Certains processus théoriques (mais non réels) se comportent de façon "étonnante".

Prenons en effet :

---

(24) On vérifie bien ici que M_1 et M_2 ayant *toutes les 2* le vecteur P_0 comme vecteur stationnaire appartiennent *toutes les 2* au segment défini par :

$$M = \lambda \begin{pmatrix} .6 & .4 \\ .6 & .4 \end{pmatrix} + \mu \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{avec} \quad \lambda + \mu = 1\ , (\lambda\ , \mu) \geqslant (0\ , 0)$$

$$M_1 = \begin{pmatrix} .96 & .04 \\ .18 & .82 \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} .7 & .3 \\ 1.1 & -.1 \end{pmatrix}$$

On en déduit :

$$M_2 = C\,M_1 = \begin{pmatrix} .7 & .3 \\ 1.1 & -.1 \end{pmatrix} \begin{pmatrix} .96 & .04 \\ .18 & .82 \end{pmatrix} = \begin{pmatrix} .726 & .274 \\ 1.038 & -.038 \end{pmatrix}$$

(C et M_2 ne sont pas stochastiques) et

$$\epsilon_1 = -.234$$
$$\epsilon_2 = .858$$
$$k = -2.5$$
$$\lambda = \frac{1}{k} = -.4$$

Cependant, d'après Harary-Lipstein-Smith, théorème 2[25], λ étant compris strictement entre -1 et 0, non seulement la suite $M_1, M_2, \ldots M_n, \ldots$ devrait converger, mais tous ses éléments devraient être stochastiques, ce qui est *manifestement faux.*

Dans leur article ils considèrent une matrice de causalité à droite (c'est-à-dire $M_2 = M_1 K$) alors que nous avons choisi une matrice de causalité à gauche (c'est-à-dire $M_2 = C\,M_1$), mais comme C et K sont 2 *matrices "semblables"*[26] cela ne change rien au résultat.

Toujours à propos de ce *contre exemple,* il est bon de noter qu'après ce comportement *"erratique",* les matrices de *"transition"* se comportent de façon *"plus sage",* puisque l'on aurait :

$$M_3 = \begin{pmatrix} .8196 & .1804 \\ .6948 & .3052 \end{pmatrix}, \quad M_4 = \begin{pmatrix} .78216 & .21784 \\ .83208 & .16792 \end{pmatrix}$$

$$M_5 = \begin{pmatrix} .797136 & .202864 \\ .777168 & .222832 \end{pmatrix} \ldots,$$ et que l'on a M_n stochastique

pour $n \geqslant 3$ et $M_n \xrightarrow[n \to \infty]{} \begin{pmatrix} .7928572, \ldots \\ .7928572, \ldots \end{pmatrix}$

c'est dire que, *à partir de la période 3,* le processus envisagé peut se produire *dans la pratique.*

---

(25) "Let λ the non unitcharacteristic root of the causative matrix C. Whenever

$$0 \leqslant \lambda < 1, \Pi_{s=0}^{s=\infty} (MC^s)$$

is stochastic if andonly if C is. *Whenever* $-1 < \lambda \leqslant 0$, MC^s *is stochastic for all* $s = 0, 1, \ldots,$ and so are $\Pi_{s=0}^{s=2} (MC^s)$ and $\Pi_{s=0}^{s=\infty} (MC^s \ldots$"

(26) au sens algébrique du terme.

IV – DETERMINATION DU PARAMETRE k DANS LE CADRE D'UNE MATRICE CAUSALE D'ORDRE 2, SEULES LES PARTS DE MARCHE ETANT SUPPOSEES CONNUES

Nous poserons $E_1 = E = \begin{pmatrix} \epsilon_1 & -\epsilon_1 \\ \epsilon_2 & -\epsilon_2 \end{pmatrix}$

et

$$M_1 = \begin{pmatrix} m_1 & 1 - m_1 \\ p_1 & 1 - p_1 \end{pmatrix} \quad , \quad M_2 = \begin{pmatrix} m_2 & 1 - m_2 \\ p_2 & 1 - p_2 \end{pmatrix}$$

Rappelons que nous avions :

$$k = 1 + \frac{\epsilon_2 - \epsilon_1}{m_2 - p_2} = \frac{m_1 - p_1}{m_2 - p_2} \qquad (= d_1 - d_2)$$

$$\lambda = \frac{1}{k} = \frac{m_2 - p_2}{m_1 - p_1} = 1 + \frac{\epsilon_2 - \epsilon_1}{m_1 - p_1} \tag{1}$$

Nous allons montrer que : λ *est solution d'une équation algébrique du 3ème degré,* les coefficients étant fonction *uniquement des parts de marché aux 5 premières périodes.*

(1) peut s'écrire

$$\boxed{0 = (1 - \lambda)(m_1 - p_1) + (\epsilon_1 - \epsilon_2)} \tag{1bis}$$

En posant $P_n = (a_n \, , \, 1 - a_n) = P_{n-1} M_n = P_{n-1} \left(M_{n-1} + \frac{E}{k^{n-2}} \right)$

nous avons

(1ter) $\quad a_1 = (m_1 - p_1)\, a_0 + p_1$

(27) $$\boxed{a_2 - a_1 = (m_1 - p_1)(a_1 - a_0) + (\epsilon_1 - \epsilon_2)\, a_1 + \epsilon_2} \tag{2}$$

(27) $$\boxed{a_3 - a_1 = (m_1 - p_1)(a_2 - a_0) + (\epsilon_1 - \epsilon_2)\, a_2 (1 + \epsilon_2 (1 + \lambda)} \tag{3}$$

(27) $$\boxed{a_4 - a_1 = (m_1 - p_1)(a_3 - a_0) + (\epsilon_1 - \epsilon_2)\, a_3 (1 + \lambda + \lambda^2)}$$
$$\boxed{+ \epsilon_2 (1 + \lambda + \lambda^2)} \tag{4}$$

- - - - - - - - - - - - - -

(27) On aurait, de même :

$$a_2 - a_2 = (m_2 - p_2)(a_2 - a_1) + \lambda(\epsilon_1 - \epsilon_2)\, a_2 + \lambda \epsilon_2$$
$$= \lambda(m_1 - p_1)(a_2 - a_1) + (\epsilon_1 - \epsilon_2)\, a_2 + \epsilon_2)$$
$$a_4 - a_2 = \lambda[(m_1 - p_1)(a_3 - a_1) + (\epsilon_1 - \epsilon_2)\, a_3 (1 + \lambda) + \epsilon_2 (1 + \lambda)]$$
$$a_5 - a_2 = \lambda[(m_1 - p_1)(a_4 - a_1) + (\epsilon_1 - \epsilon_2)\, a_4 (1 + \lambda + \lambda^2) + \epsilon_2 (1 + \lambda + \lambda^2)]$$

Le système de 4 équations (1bis, 2, 3, 4) à 3 inconnues "$m_1 - p_1$" "$\epsilon_1 - \epsilon_2$", "ϵ_2" devant être compatible, le déterminant de ce système doit être nul ; on a donc :

$$\begin{vmatrix} \lambda = 1 & 1 & 0 & 0 \\ -(a_1 - a_0) & +a_1 & +1 & a_2 - a_1 \\ -(a_2 - a_0) & +a_2(1+\lambda) & (1+\lambda) & a_3 - a_2 \\ -(a_3 - a_0) & a_3(1+\lambda+\lambda^2) & (1+\lambda+\lambda^2) & a_4\ a_1 \end{vmatrix} = 0$$

ou, en ajoutant à la 1ère colonne la 2ème colonne et l'opposée de la 3ème colonne, pondérée par a_0 :

$$\begin{vmatrix} \lambda & 1 & 0 & 0 \\ 0 & a_1 & 1 & a_2 - a_1 \\ -\lambda(a_2 - a_0) & a_2(1+\lambda) & (1+\lambda) & a_3 - a_1 \\ -\lambda - \lambda^2(a_3 - a_0) & a_3(1+\lambda+\lambda^2) & (1+\lambda+\lambda^2) & a_4 - a_1 \end{vmatrix} = 0$$

ce qui donne finalement :

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & a_1 - a_0 & 1 & a_2 - a_1 \\ a_2 - a_0 & \lambda(a_2 - a_0) & 1+\lambda & a_3 - a_1 \\ a_3 - a_0 & \lambda^2(a_3 - a_0) & (1+\lambda+\lambda^2) & a_4 - a_1 \end{vmatrix} = 0$$

ou, en développant :

$$\boxed{\begin{aligned} &\lambda^3[(a_2 - a_1)(a_2 - a_3)] + [(a_2 - a_1)(a_2 - a_3) + (a_3 - a_1)^2] \\ &+ \lambda[(a_1 - a_0)(a_4 - a_3) + (a_2 - a_0)(a_2 - a_4)] + (a_1 - a_0)(a_4 - a_3) = 0 \end{aligned}}$$

c.q.f.d.

On peut même voir que $\lambda = \frac{1}{k}$ *est la seule racine commune* aux 2 polynomes du 3ème degré f et g donnés par :

$$f(X) = Z_3X^3 + Z_2X^2 + Z_1X^1 + Z_0$$

et

$$g(X) = Y_3X^3 + Y_2X^2 + Y_1X^1 + Y_0$$

$Z_3 = (a_2 - a_1)(a_2 - a_3)$	$Y_3 = (a_3 - a_2)(a_3 - a_4)$
$Z_2 = (a_2 - a_1)(a_2 - a_3) + (a_3 - a_1)^2$	$Y_2 = Y_3 + (a_4 - a_2)^2$
$Z_1 = (a_1 - a_0)(a_4 - a_3) + (a_2 - a_0)(a_2 - a_4)$	$Y_1 = Y_0 + (a_3 - a_1)(a_3 - a_5)$
$Z_0 = (a_1 - a_0)(a_4 - a_3)$	$Y_0 = (a_2 - a_1)(a_5 - a_4)$

Ces polynomes ont une seule racine commune, donnée par la formule[28].

$$\alpha = -\frac{\begin{vmatrix} Z_3 & 0 & 0 & Y_3 \\ Z_2 & Z_3 & Y_3 & Y_2 \\ Z_1 & Z_2 & Y_2 & Y_1 \\ 0 & Z_0 & Y_0 & 0 \end{vmatrix}}{\begin{vmatrix} Z_3 & 0 & 0 & Y_3 \\ Z_2 & Z_3 & Y_3 & Y_2 \\ Z_1 & Z_2 & Y_2 & Y_1 \\ Z_0 & Z_1 & Y_1 & Y_0 \end{vmatrix}} = -\frac{D_1}{\Delta_1}$$

ou, en développant :

$$D_1 = (Z_2 Y_0 - Y_2 Y_0)(Z_3 Y_2 - Y_3 Z_2) + (Z_3 Y_0 - Y_3 Z_0)(Z_1 Y_3 - Z_3 Y_1)$$

$$\Delta_1 = Z_3^2 (Y_2 Y_0 - Y_1 Y_1) - Z_2 Y_3 (Z_2 Y_0 - Y_1 Z_1) + Z_3 Y_2 (Z_2 Y_1 - Y_2 Z_1)$$
$$- Y_3 Z_2 (Z_2 (Z_2 Y_1 - Y_2 Z_1) + Z_2 Y_3 (Z_1 Y_1 - Y_2 Z_0) - Y_3^2 (Z_1 Z_1 - Z_2 Z_0)$$

c'est-à-dire que : $\lambda = 1/k$ *s'obtient de façon unique si on connaît les parts de marché des 6 premières périodes.*

V – DETERMINATION DU PARAMETRE k DANS LE CADRE D'UNE MATRICE CAUSALE D'ORDRE 2, LES PARTS DE MARCHE ET LA PREMIERE MATRICE DE TRANSITION ETANT SUPPOSEES CONNUES.

Nous avons, comme précédemment, les équations :

---

(28) Voir, par exemple, le cours d'algèbre de Lelong-Arnaudies, chez Dunod.

(29) $$0 = (1 - \lambda)(m_1 - p_1) + \epsilon_1 - \epsilon_2 \qquad (1)$$

(29) $$a_2 - a_1 = (m_1 - p_1)(a_1 - a_0) + (\epsilon_1 - \epsilon_2)\, a_1 + \epsilon_2 \qquad (2)$$

(29) $$a_3 - a_1 = (m_1 - p_1)(a_2 - a_0) + (\epsilon_1 - \epsilon_2)\, a_2 (1 + \lambda) + \epsilon_2 (1 + \lambda) \qquad (3)$$

Ce système à 3 équations à 2 inconnues "$\epsilon_1 - \epsilon_2$" et "ϵ_2" devant être compatible, le déterminant de ce système doit être nul ; on a donc :

$$\begin{vmatrix} 1 & 0 & (\lambda - 1)(m_1 - p_1) \\ a_1 & 1 & (a_2 - a_1) - (a_1 - a_0)(m_1 - p_1) \\ a_2(1 + \lambda) & 1 + \lambda & (a_3 - a_1) - (a_2 - a_0)(m_1 - p_1) \end{vmatrix} = 0$$

ce qui donne

$$\begin{vmatrix} 1 & 0 & \lambda(m_1 - p_1) \\ a_1 & 1 & a_2 - a_1 \\ a_2(1 + \lambda) & 1 + \lambda & (a_3 - a_1) + (a_2 - a_0)\lambda(m_1 - p_1) \end{vmatrix} = 0$$

ou encore

$$\begin{vmatrix} 1 & 0 & -\lambda(m_1 - p_1) \\ a_2 - a_1 & -1 & a_2 - a_1 \\ 0 & -(1 + \lambda) & (a_3 - a_1) + (a_2 - a_0)\lambda(m_1 - p_1) \end{vmatrix} = 0$$

ce qui donne, en développant et en ordonnant par rapport à λ

$$\boxed{\lambda^2 [(m_1 - p_1)(a_3 - a_1)] + \lambda[(m_1 - p_1)(a_0 - a_1) + (a_2 - a_1)] + (a_2 - a_3) = 0}$$

Ainsi, nous avons :

Théorème : $\lambda = \dfrac{1}{k} = \dfrac{m_2 - p_2}{m_1 - p_1}$ est solution d'une équation algébrique du 2ème degré, les coefficients étant fonction uniquement de "$m_1 - p_1$" et des parts de marché aux 4 premières périodes

En résumé :

1/ *La seule connaissance des parts de marché (ou vecteur d'états) des 6 premières périodes suffit pour déterminer tous les paramètres du modèle à causalité constante*(30).

---

(29) Et nous aurions, de même :

(2bis) $a_3 - a_2 = \lambda[(m_1 - p_1)(a_2 - a_1) + (\epsilon_1 - \epsilon_2)\, a_2 + \epsilon_2)$

(3bis) $a_4 - a_2 = \lambda[(m_1 - p_1)(a_3 - a_1) + (\epsilon_1 - \epsilon_2)\, a_3(1 + \lambda) + \epsilon_2(1 + \lambda)]$

(30) En fait, les auteurs ont mis au point un programme informatique utilisant le théorème du point fixe qui permet d'avoir besoin uniquement des 5 premières parts de marché.

2/ *La seule connaissance des parts de marché des 5 premières périodes et de la première matrice de transition suffit pour déterminer tous les paramètres du modèle.*

Remarque : Les algorithmes précédents sont de type *analytique,* c'est-à-dire qu'il n'est pas ici tenu compte des fluctuations *aléatoires non expliquées par le modèle.* Pour cerner de plus près la réalité des données nous avons donc mis au point certains programmes informatiques, dont, en particulier, un programme de régression qui, tant sur des données simulées que sur des données réelles, donnent des résultats assez riches. Il reste encore beaucoup à faire, cependant, en ce domaine tant dans le cadre de la manipulation des données[31] que dans le cadre des tests statistiques proprements dits (Test d'un processus homogène contre un processus à causalité constante, test sur les paramètres du modèle, . . .).

Nous croyons de plus, tant en raison de la relative simplicité de notre modèle qu'en raison du fait qu'il généralise tous les résultats utilisés jusqu'alors dans ce domaine, que ce nouveau modèle peut contribuer de façon efficace à l'analyse statistique du comportement du consommateur dans le domaine des biens de grande consommation. Ce type de modèle, à comportement asymptotique "géométrique" paraît, enfin, rendre compte de l'effet de la publicité dans l'évolution des parts de marché.

---

(31) Un algorithme de construction de matrices de transition, de type nouveau, vient d'être mis au point par les auteurs.

BIBLIOGRAPHIE

Articles :

(1) T.W. ANDERSON and LEO. A. GOODMAN – Statistical Inference about Markov Chaîns – *Annals of Mathematics Statistics 28 (1957).*

(2) W. DENT – A Note on Lipstein's Model of Consumer Behavior *(Operations Research, 1973).*

(3) M.M. DRYDEN – Share Price Movements : a Markovian Approach *(Journal of Finance, 24, 1969).*

(4) A.S.C. EHRENBERG – An Appraisal of Markov Brand-Switching Models – *Journal of Marketing Research, Vol. II, Nov. 1965).*

(5) R.E. FRANCK – Brand and Choice as a Probability Process *Journal of Business 35 n° 1 (janvier 1962 – pages 43 à 56).*

(6) F. HARARY, B. LIPSTEIN – The Dynamics of Brand Loyalty : a Markovian Approach – *(Operations Reserach 10, Feb. 1962).*

(7) F. HARARY, B. LIPSTEIN, G.P.H. STYAN – A Matrix Approach to Nonstationary Chaîns – *(Operations Reserach, N-D, 70).*

(8) T.C. LEE, G.C. JUDGE, R.L. CAIN – A Sampling Study of the Properties of Estimators of Transition Probabilities *(Management Science, Vol. 15, n° 7, March 1969).*

(9) Bruno LE MAIRE, Gilles MAUFFREY – Nouvelle approche Markovienne des Phénomènes de "Brand Loyalty and Brand-Switching" *(Les Cahiers de Recherche – C.E.S.A.).*

(10) B. LIPSTEIN – A Mathematical Model of Consumer Behavior *(Journal of Marketing Research, 2, 1965).*

(11) W.F. MASSY – Order and Homogeneity of Family Specific Brand-Switching Processes *(Journal of Marketing Research, Vol. III, Jan. 1966).*

(12) W.F. MASSY, D.G. MORRISON – Comments on Ehrenberg's Appraisal of Brand-Switching Models – *(Journal of Marketing Research May 1968).*

Livres :

(1) BARTHOLOMEW – Stochastics Models for Social Processes – *WILEY*

(2) COX and MILLER – The Theory of Stochastic Processes – *CHAPMAN and HALL.*

(3) EHRENBERG – Repeat Buying – *NORTH HOLLAND.*

(4) KEMENY-SNELL – The Theory of Stochastic Processes – *CHAPMAN and HALL*

(5) MASSY, MONTGOMERY, MORRISON – Stochastics models of buying behavior – *M.I.T. PRESS.*