Université de Nice — Année 2011-2012
Nice-Polytech — MMDFA

Cours 11 et 12 : chaînes de Markov et théorème de Perron-Frobenius appliqués à la micro-finance

On modélise avec des chaînes de Markov l'évolution au cours du temps d'une quantité X qui peut prendre un nombre fini d'états $X = x_1$, $X = x_2$, …, $X = x_n$ et qui passe de l'état x_i à l'instant t à l'état x_j à l'instant suivant $t+1$ avec une probabilité p_{ij} donnée. Les nombres $p_{ij} = P(X_{t+1} = x_j / X_t = x_i)$ vérifient donc $0 \leq p_{ij} \leq 1$ et $\sum_{j=0}^{n} p_{ij} = 1$ (puisque si la chaîne est dans l'état x_i à un instant, elle sera nécessairement dans l'un des états possibles $x_1, \ldots, x_n$ l'instant suivant et donc $p_{i1} + p_{i2} + \ldots + p_{in} = 1$).

Pour définir une chaîne de Markov il faut donc deux ingrédients de base :

1. L'*espace des états* $S := \{x_1, \ldots, x_n\}$ connu que l'on supposera fini
2. La *matrice de transition* (ou de passage)

$$\mathbb{P} = (p_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} = \begin{array}{cc} \begin{array}{cccc} x_1 & x_2 & \ldots & x_n \end{array} & \\ \left(\begin{array}{cccc} p_{11} & p_{12} & \cdots & p_{1n} \\ \vdots & & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{array} \right) & \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \end{array}$$

A noter que cette matrice est appelée *matrice stochastique* parce que ses coefficients sont tous compris entre 0 et 1 et la somme des coefficients de chaque ligne vaut 1 (mais ce n'est pas vrai en général pour les colonnes).

On peut aussi représenter une chaîne de Markov $(S, \mathbb{P})$ par un *diagramme en points et flèches* comme indiqué par la figure (1) correspondant à l'exemple ci-dessous. Dans ces diagrammes, chaque état est représenté par un point et chaque coefficient p_{ij} non nul de la matrice de transition par une flèche allant de l'état i à l'état j.

1 Un exemple en microfinance

Le modèle suivant est du à une économiste américaine, G. Tedeschi, et il étudie le méchanisme du microcrédit dans une population d'emprunteurs. Dans sa version la plus simple, les individus peuvent être dans deux états, l'état de demandeur (d'un prêt) et celui de bénéficiaire (d'un prêt). On suppose que la seule sanction en cas de non remboursement est la perte du droit automatique à un nouveau prêt, droit qui, au contraire, est garanti au bénéficiaire lorsqu'il rembourse son prêt. L'objectif est d'étudier l'effet de cette incitation (la garantie d'une reconduction automatique du prêt) sur le taux de remboursement et la répartition entre demandeurs et bénéficiaires au sein de la population. On suppose qu'un bénéficiaire qui investi l'argent emprunté dans une micro entreprise sera en mesure de rembourser son emprunt (et choisira de le rembourser effectivement) avec une probabilité α alors qu'il fera defaut avec une probabilité $1 - \alpha$. En cas de defaut, il redevient simplement demandeur durant la période suivante. Comme demandeur, on suppose qu'il a une probabilité γ de se voir attribuer un prêt et une probabilité $1 - \gamma$ de se le voir refusé.

On introduit donc une chaîne de Markov $(X_t)_{t \geq 0}$ à deux états $S = \{B, D\}$ et dont la matrice de passage est

$$\mathbb{P} = \begin{array}{cc} \begin{array}{cc} B & D \end{array} & \\ \left(\begin{array}{cc} \alpha & 1-\alpha \\ \gamma & 1-\gamma \end{array} \right) & \begin{array}{c} B \\ D \end{array} \end{array}$$

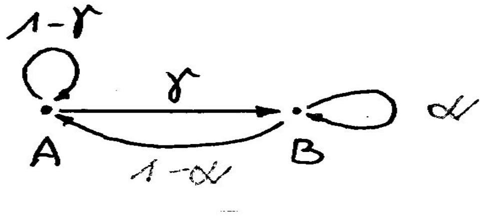

FIG. 1 – Diagramme en points et flèches correspondant à l'exemple de la dynamique de Tedeschi sans période d'exclusion.

d'où le diagramme en points et flèches de la figure (1).

On peut ainsi calculer la probabilité de n'importe quelle succession d'états, appelée *trajectoire* de la chaîne de Markov. Par exemple la probabilité qu'un bénéficiaire à la première période reste bénéficiaire pendant 2 périodes puis revienne demandeur et le reste pendant 2 périodes avant de redevenir bénéficiaire, c'est-à-dire la probabilité de la succession d'états (B, B, D, D, B) est égale à

$$\begin{aligned} P \quad & (X_0 = B, X_1 = B, X_2 = D, X_3 = D, X_4 = B) \\ = \quad & \pi_0(B)P(X_1 = B/X_0 = B)P(X_2 = D/X_1 = B)P(X_3 = D/X_2 = D)P(X_4 = B/X_3 = D) \\ = \quad & \pi_0(B)p_{11}p_{12}p_{22}p_{21} = \pi_0(h)\alpha(1-\alpha)(1-\gamma)(\gamma) = \alpha\gamma(1-\alpha)(1-\gamma)\pi_0(B). \end{aligned}$$

2 Evolution de la distribution initiale

Mais on ne cherche pas seulement à calculer la probabilité particulière de chaque trajectoire de notre chaîne de Markov, on voudrait plus généralement déterminer l'évolution des proportions des différents états entre le premier et le deuxième instant, entre le deuxième et le troisième, et plus généralement savoir comment vont évoluer ces proportions à l'avenir.

Si l'on connait la distribution initiale des différents états (c'est-à-dire la proportion d'individus de la population étudiée se trouvant dans chacun des états x_i, que l'on appelle la *loi de probabilité initiale* π_0), l'étude de la chaîne de Markov permet de calculer, à partir de cette répartition

S	x_1	x_2	....	x_n
π_0	$\pi_0(x_1)$	$\pi_0(x_2)$	....	$\pi_0(x_n)$

$\pi_0(x_i)$ à l'instant $t = 0$, c'est-à-dire à partir des nombres $\pi_0(x_i) := P(X_0 = x_i)$, quels états la population va atteindre à l'instant $t = 1$ et avec quelles probabilités π_1, puis à l'instant $t = 2$ et ainsi de suite. En d'autres termes, on va ainsi calculer la loi π_t pour tous les $t > 0$ et ainsi modéliser la dynamique de cette population.

Voici comment on procède dans l'exemple considéré. Pour calculer les probabilités π_1 des deux états à l'instant $t = 1$, c'est-à-dire pour calculer

$$\pi_1 := (P(X_1 = B), P(X_1 = D)) = (\pi_1(B), \pi_1(D)),$$

on remarque que, du fait de la formule des probabilités totales, $\pi_1(B) = P(X_1 = B/X_0 = B)P(X_0 = B) + P(X_1 = B/X_0 = D)P(X_0 = D)$, ce qui peut s'écrire ici $\pi_1(B) = \alpha\pi_0(B) + \gamma\pi_0(D)$ compte tenu des valeurs des probabilités de transition données par la matrice $\mathbb{P}$. De même, on vérifie facilement que $\pi_1(D)$ est le produit scalaire du vecteur π_0 avec la deuxième colonne de la matrice $\mathbb{P}$. Donc le vecteur π_1 est le produit du vecteur π_0 par la matrice $\mathbb{P}$, ce qui s'écrit simplement $\pi_1 = \pi_0\mathbb{P}$, donc

$$(\pi_1(B), \pi_1(D)) = (\pi_0(B), \pi_0(D)) \begin{pmatrix} \alpha & 1-\alpha \\ \gamma & 1-\gamma \end{pmatrix}.$$

Pour une chaîne de Markov d'espace d'états S et de matrice de transition $\mathbb{P}$, l'évolution au cours du temps de la loi de probabilité initiale π_0 est donnée par $\pi_1 = \pi_0\mathbb{P}$, $\pi_2 = \pi_1\mathbb{P} = (\pi_0\mathbb{P})\mathbb{P} = \pi_0\mathbb{P}^2$,... et plus généralement $\pi_t = \pi_{t-1}\mathbb{P} = \ldots = \pi_0\mathbb{P}^n$. En particulier, la matrice de transition pour passer de l'état t à l'état $t+2$ est égale à $\mathbb{P} \times \mathbb{P} = \mathbb{P}^2$, et plus généralement, celle pour passer de l'état t à l'état $t+k$ est égale à $\mathbb{P}^k$. C'est une caractéristique importante des chaînes de Markov que la matrice de transition $\mathbb{P}$ élevée à la puissance k contient les probabilités de transitions de chacun des états vers les autres en exactement k intervalles de temps.

Définition : Une loi de probabilité π sur l'espace des états S est appelée *stationnaire* pour la chaîne de Markov $(S, \mathbb{P})$ si la chaîne laisse la loi inchangée, c'est-à-dire si l'on a $\pi\mathbb{P} = \pi$.

La *distribution stationnaire d'une chaine de Markov* est donc une répartition du système entre ses divers états ayant la propriété de rester inchangée au cours du temps. Plus précisément, si les proportions des différents états à l'instant initial sont celles d'une distribution stationnaire, alors ces proportions ne sont plus modifiées ultérieurement par la dynamique de la chaîne de Markov. On peut donc comprendre les distributions stationnaires comme une sorte d'équilibre pour le système. Une chaîne de Markov peut ne posséder aucun équilibre de la sorte, ou bien en posséder un ou plusieurs. Mais ces équilibres, lorsqu'ils existent, n'ont pas tous la même pertinence. Ceux qui vont compter vraiment sont les équilibres vers lesquels on se rapproche inéluctablement lorsque le temps s'écoule, et ce quelque soit la distribution initiale du système. Or si l'on observe les puissances *grandes* de la matrice de transition $\mathbb{P}$ (ce que l'on peut faire facilement avec Scilab par exemple), on peut voir qu'en général, lorsque k augmente, la matrice $\mathbb{P}^k$ se stabilise peu à peu et prend la forme d'une matrice dont toutes les lignes sont (presque) égales.

On peut alors faire la remarque générale suivante : une matrice stochastique (donc nécessairement carrée) $M = \begin{pmatrix} L_1 \\ ... \\ L_n \end{pmatrix}$ dont toutes les lignes sont égales à un même vecteur ligne L^* vérifie l'égalité $L^*M = L^*$. Comme L^* est de plus un vecteur dont la somme des coefficients vaut 1 (comme ligne d'une matrice stochastique), cette propriété fait de L^* un bon candidat pour être la distribution stationnaire que l'on recherche c'est-à-dire la distribution d'équilibre vers laquelle le système évolue. Nous verrons pourquoi c'est effectivement la distribution stationnaire recherchée lors du prochain cours.

3 Etat stationnaire du modèle de Tedeschi

Considérons la dynamique suivante qui modélise l'évolution des proportions de *bénéficiaires* (B) et de *demandeurs* (D) dans une population d'individus susceptibles de bénéficier d'une activité de micro crédit, $S, \mathbb{P}$ où $S = \{B, D\}$ et

$$\mathbb{P} = \begin{array}{c} \begin{array}{cc} B & D \end{array} \\ \begin{pmatrix} \alpha & 1-\alpha \\ \gamma & 1-\gamma \end{pmatrix} \end{array} \begin{array}{c} \\ B \\ D \end{array}$$

La recherche d'un vecteur propre à gauche $\pi*$ de valeur propre 1 (solution de $\pi\mathbb{P} = \pi$) donne facilement le vecteur $(\gamma, 1-\alpha)$ (ou tout autre multiple de ce vecteur) mais ce dernier n'est pas une distribution stationnaire, sauf si $\gamma + 1 - \alpha = 1$. Il convient donc de diviser ses deux composantes par la quantité $\gamma + 1 - \alpha$ pour obtenir la distribution stationnaire $\pi*$ recherchée. Par exemple si la probabilité de réussite des bénéficiaires est de 75% et la chance d'obtenir un prêt pour le demandeur de 50%, c'est-à-dire si $\alpha = 0.75$ et $\gamma = 0.5$, alors on trouve $\pi* = (\frac{2}{3}, \frac{1}{3})$. A présent, si l'on observe ce qui se passe si l'on part d'une distribution initiale π_0 différente de la distribution stationnaire, les itérés successifs π_0, $\pi_1 = \pi_0\mathbb{P}$, $\pi_2 = \pi_1\mathbb{P} = \pi_0\mathbb{P}^2$, ... (que l'on peut facilement calculer avec Scilab par exemple), on peut voir que, lorsque k augmente, la distribution $\pi_0\mathbb{P}^k$ se stabilise peu à peu et converge vers $\pi*$. Par exemple s'il y a au départ 30% de bénéficiaires pour 70% de demandeurs, $\pi_0 = (0.3, 0.7)$, on obtient $\pi_1 = (0.575, 0.425)$, $\pi_2 = (0.644, 0.356)$, $\pi_3 = (0.661, 0.339)$, et $\pi_4 = (0.665, 0.335)$, etc...Cela signifie, dans cet exemple, que quelque soit la répartition initiale des bénéficiaires et des demandeurs, l'activité va faire évoluer très vite cette proportion vers l'équilibre $(\frac{2}{3}, \frac{1}{3})$. Cette répartition limite dépend des deux probabilités α et γ et on peut vérifier facilement que la proportion de bénéficiaires est d'autant plus grande que ces deux paramètres sont grands. Pour l'IMF, augmenter α peut consister à n'accorder de prêt que lorsqu'elle peut s'assurer que le projet à financer a de très bonnes chances de réussir mais cela se fera alors probablement au détriment de γ car il lui faudrait alors rejeter une plus grande proportion de demandes.

A noter enfin que, si l'on calcule une puissance grande de la matrice $\mathbb{P}$, par exemple $\mathbb{P}^{50}$, on trouve

$$\mathbb{P}^{50} = \begin{pmatrix} 0,666667 & 0,333333 \\ 0,666667 & 0,333333 \end{pmatrix}.$$

Les deux lignes de la matrices sont (presque) identiques et (presque) égales à la distribution stationnaire $\pi*$. On peut le comprendre de la façon suivante. Si l'on a bien, comme nous le verrons ci-dessous, $\lim_{k\to+\infty} \pi_0\mathbb{P}^k = \pi*$ pour n'importe quelle distribution initiale π_0, le $k$ième itéré de π_0, π_k, est proche de $\pi*$ dès que k est suffisamment grand. Comme $\pi_k = \pi_0\mathbb{P}^k \simeq \pi*$, si la matrice $\mathbb{P}^k$ a ses deux lignes égales à L, alors on vérifie facilement que $\pi_0\mathbb{P}^k = L$ et donc que $L \simeq \pi*$. Comme L est de plus un vecteur dont la somme des coefficients vaut 1 (comme ligne d'une matrice stochastique), cette propriété fait de L un bonne approximation de la distribution stationnaire que l'on recherche, c'est-à-dire la distribution d'équilibre vers laquelle le système évolue. En pratique, les lignes de $\mathbb{P}^{40}$ ne sont pas, en général, exactement identiques, donc L n'est en réalité qu'une approximation de la répartition d'équilibre.

4 La théorie de Perron-Frobenius

Cette théorie mathématique qui porte le nom de ses deux inventeurs allemands, Ferdinand Georg Frobenius (1949-1917) et Oskar Perron (1880-1975), concerne les matrice positives dont l'une des puissances est strictement positive. Ces matrices s'appellent des matrices *primitives*.

Définition : On dit qu'une matrice carrée M est *positive* si tous ses coefficients sont positifs ou nuls et *strictement positive* si elle est positive et n'a aucun coefficient nul. Une matrice positive M est dite *primitive* s'il existe un entier $k \geq 0$ tel que M^k est strictement positive.

Théorème 1 *Toute matrice stochastique M qui est primitive possède un unique vecteur propre stochastique strictement positif, $\pi*$, de valeur propre 1 qu'on appelle* vecteur propre dominant de Perron-Frobenius[1] *et, pour tout vecteur π_0 stochastique, on a*

$$\lim_{k \to +\infty} \pi_0 M^k = \pi * .$$

Appliqué au cas d'une chaîne de Markov, ce résultat assure que si la matrice de transition de la chaîne est primitive, alors, quelque soit la distribution initiale π_0 du système, la répartition des états va évoluer, sous l'action de la dynamique de la chaîne de Markov, vers une distribution stationnaire $\pi*$ qui constitue un équilibre pour le système.

En fait, la théorie de Perron-Frobenius ne s'applique pas uniquement aux matrices stochastiques, mais plus généralement à toutes les matrice positives (pourvu qu'elles soient *irréductibles*, ce qui est une propriété un peu plus générale que celle d'être primitive.

L'une des difficultés de son application est la vérification de son hypothèse principale, la propriété pour la matrice d'être primitive. En effet, si l'on trouve, en calculant des puissances de la matrice, l'une d'entre elles qui est strictement positive, on saura que la matrice est primitive. Mais lorsque toutes les puissances calculées comportent des coefficients nuls, on ne sait pas alors si la matrice est primitive ou non Une *recette* peut cependant être utile :
Une matrice $\mathbb{P}$ de taille $n \times n$ est primitive si et seulement si $\mathbb{P}^{n^2-2n+2} > 0$.

En pratique, si l'on veut déterminer l'évolution future d'un système modélisé par une chaîne de Markov, on procèdera de la façon suivante :

1. Etablir (si c'est bien le cas) que la matrice de transition est une matrice primitive.
2. A l'aide d'un logiciel de calcul scientifique (par exemple Scilab), calculer une puissance grande de la matrice de transition, par exemple $\mathbb{P}^{40}$ ou $\mathbb{P}^{100}$.
3. Cette puissance ayant toutes ses lignes (presque) égales, cette ligne est la distribution stationnaire recherchée.
4. La théorie de Perron-Frobenius permet alors d'affirmer que la dynamique considérée fait évoluer le système, quelque soit la distribution initiale, vers l'équilibre donné par ce vecteur propre.

Remarque : Pour donner l'intuition de ce qui se passe pour la dynamique d'une chaine de Markov lorsque sa matrice de transition n'est pas primitive, on peut envisager le cas d'une chaîne *périodique*. C'est le cas d'une chaîne ayant une matrice de transition dont l'une des puissances $\mathbb{P}^k$, $k = 2, 3, \ldots$, vérifie $\mathbb{P}^k = \mathbb{P}$. Par exemple on pourra vérifier que pour

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

on a $\mathbb{P}^4 = \mathbb{P}$; cette chaîne est dite *périodique de période* 3 car toute trajectoire revient à son état initial après 3 étapes. Il ne peut donc pas y avoir d'évolution vers un état stationnaire pour une telle chaine de Markov.

[1] Le *Pagerank* utilisé par le moteur de recherche Google est un exemple de vecteur propre dominant de Perron-Frobenius (voir `http ://fr.wikipedia.org/wiki/Théorème_de_Perron-Frobenius`).