Université de Nice — L1-SV, année 2011-2012
Département de Mathématiques — Mathématiques Appliquées à la Biologie (semestre1)

Cours 1 : Introduction aux chaînes de Markov

Pour modéliser l'évolution au cours du temps (la dynamique) de systèmes biologiques, par exemple celle d'un organisme, d'une substance, d'un écosystème, on choisit souvent des *modèles aléatoires.* Les plus simples de ces modèles aléatoires sont les chaînes de Markov qui, dans le cas particulier étudié ici sont faciles à utiliser car il n'y a pratiquement aucun prérequis mathématique. Elles nous donneront l'occasion d'une première familiarisation avec le calcul matriciel que nous approfondirons lors des leçons suivantes.

1 Un exemple en écologie

On s'intéresse au développement d'une forêt naturelle en région tempérée sur une parcelle en friche (par exemple par abandon d'une zône cultivée ou suite à un incendie). Notre modèle simplifié comporte 3 états. L'état 1 est celui d'une végétation constituée d'herbes ou d'autres espèces pionnières ; l'état 2 correspond à la présence d'arbustes dont le développement rapide nécessite un ensoleillement maximal et l'état 3 celui d'arbres plus gros qui peuvent se développer dans un environnement semi ensoleillé. Si l'on note respectivement h, a, f ces trois états (pour *herbe, arbustes, forêt*), l'ensemble des états possibles pour un point donné de cette parcelle est l'ensemble $S = \{h, a, f\}$. Sur la parcelle on repère au sol un grand nombre de points (un millier) répartis sur un maillage régulier et on enregistre à intervalle de temps fixé (tous les 3 ans) l'état de la végétation en chacun de ces points (en choisissant celui des trois états qui est le plus proche).

En observant l'évolution durant un intervalle de temps, on peut déterminer pour chaque état $i \in \{h, a, f\}$ la proportion de points qui sont passés à l'état $j \in \{h, a, f\}$, et noter p_{ij} cette proportion. Si les différentes proportions ainsi relevées (il y en a 9) évoluent peu d'un intervalle de temps au suivant, on peut les supposées inchangées au cours du temps et on peut regarder comme les probabilités pour un point quelconque de passer de l'état i à l'état j pendant un intervalle de temps. Supposons par exemple que dans cette parcelle, ces probabilités soient les suivantes :

$$\mathbb{P} = \begin{array}{c} \begin{array}{ccc} h & a & f \end{array} \\ \begin{pmatrix} 0,5 & 0,45 & 0,05 \\ 0,1 & 0,5 & 0,4 \\ 0 & 0,1 & 0,9 \end{pmatrix} \end{array} \begin{array}{c} \\ h \\ a \\ f \end{array}$$

Si X_0 désigne l'état d'un point à l'instant $t = 0$ et X_1 l'état du même point en $t = 1$, on par exemple la probabilité de passage de l'état arbuste en $t = 0$ à l'état forêt en $t = 1$ s'écrit $P(X_1 = f/X_0 = a)$ est vaut $P(X_1 = f/X_0 = a) = 0,4$. De même $P(X_1 = h/X_0 = h) = 0,5$.

L'ensemble des états $S = \{h, a, f\}$ et la matrice de transition $\mathbb{P}$ constituent un exemple de *chaîne de Markov.* On peut aussi représenter cette chaine de Markov par le *diagramme en points et flèches* suivant (1).

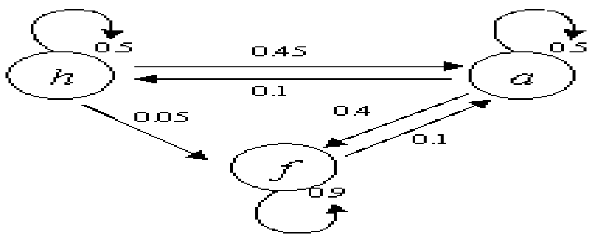

Fig. 1 – Diagramme en points et flèches correspondant à l'exemple de la dynamique de la forêt naturelle à trois états, *herbe*, *arbustes* et *forêt*, étudiée ci dessous. Notez qu'une forêt qui brûle ne devient pas de l'herbe mais est remplacée par des arbustes.

Dans ce diagramme, chaque état est représenté par une lettre encerclée et chaque coefficient non nul de la matrice de transition $\mathbb{P}$ par une flèche allant d'un état à un autre état.

Dans ce modèle, 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'en un point de la parcelle, on observe la succession d'états (h, h, a, f, f) se calcule de la façon suivante :

$$\begin{aligned} P & \quad (X_0 = h, X_1 = h, X_2 = a, X_3 = f, X_4 = f) \\ = & \quad \pi_0(h)P(X_1 = h/X_0 = h)P(X_2 = a/X_1 = h)P(X_3 = f/X_2 = a)P(X_4 = f/X_3 = f) \\ = & \quad \pi_0(h)p_{hh}p_{ha}p_{af}p_{ff} = \pi_0(h)(0,5)(0,45)(0,4)(0,9) = 0,081\pi_0(h). \end{aligned}$$

où $\pi_0(h)$ est la probabilité d'être dans l'état h à l'instant initial $t = 0$.

2 Evolution de la répartition des états au cours du temps

L'observation de l'état dans lequel se trouve les différents points de la parcelle à l'instant initial t_0 permet de déterminer les proportions initiales de chacun des 3 états $\pi_0 = (\pi_0(h), \pi_0(a), \pi_0(f))$. Pour cela, on relève pour chaque point l'état dans lequel il se trouve et on calcule la proportion de points de chacun des états possibles. On peut voir chaque proportion comme la probabilité pour un point de la parcelle d'être dans l'un des états à l'instant initial. Ainsi, si l'on a par exemple $\pi_0 = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, cela signifie que la moitié des points de la parcelle sont au départ dans l'état h, un quart dans l'état a et un quart dans l'état f. Mais on peut aussi interpréter cela en considérant qu'un état quelconque a 50% de chance d'être dans l'état h, 25% d'être dans l'état a et 25% dans l'état f. C'est pour cela que la proportion d'individus de la population étudiée se trouvant dans chacun des états,

S	h	a	f
π_0	$\pi_0(h)$	$\pi_0(a)$	$\pi_0(f)$

s'appelle la *loi de probabilité initiale* π_0 ou encore la *distribution initiale.* Lorsqu'on choisit une modélisation par une chaîne de Markov, l'objectif est souvent de déterminer l'évolution de la répartition des états au cours du temps. Par exemple, si la parcelle considérée ci-dessus est recouverte pour un tiers de forêt à l'instant initial, cette proportion va-t-elle grandir, tendre vers 100%, au contraire tendre vers zéro ou bien s'approcher d'une valeur limite, sorte d'équilibre écologique ?

Nous allons voir que si l'on connait la distribution initiale on peut calculer la distribution à l'instant $t = 1$, puis à l'instant $t = 2$ et ainsi de suite. En d'autres termes, on calcule la loi π_1

S	h	a	f
π_1	$\pi_1(h)$	$\pi_1(a)$	$\pi_1(f)$

et plus généralement la loi π_t pour tous les $t > 0$ et ainsi modéliser la dynamique de cette population.

Voici comment on procède. Pour calculer les probabilités π_1 des trois états à l'instant $t = 1$

$$\pi_1 := (P(X_1 = h), P(X_1 = a), P(X_1 = f)) = (\pi_1(h), \pi_1(a), \pi_1(f)),$$

on remarque que, du fait de la *formule des probabilités totales*[1], la probabilité $\pi_1(h)$ est égale à

$$P(X_1 = h/X_0 = h)P(X_0 = h) + P(X_1 = h/X_0 = a)P(X_0 = a) + P(X_1 = h/X_0 = f)P(X_0 = f)$$

ce qui peut s'écrire ici $\pi_1(h) = 0,5 \cdot \pi_0(h) + 0,1 \cdot \pi_0(a) + 0 \cdot \pi_0(f)$ compte tenu des valeurs des probabilités de transition données par la première colonne de la matrice $\mathbb{P}$.

On en déduit que $\pi_1(h)$ est le *produit scalaire* du vecteur π_0 avec la première colonne de la matrice $\mathbb{P}$. De même, on vérifie que $\pi_1(a)$ est le produit scalaire du vecteur π_0 avec la deuxième colonne de la matrice $\mathbb{P}$ et que $\pi_1(f)$ est le produit scalaire du vecteur π_0 avec la troixième colonne de la matrice $\mathbb{P}$. On résume cela en disant que le vecteur π_1 est le produit du vecteur π_0 par la matrice $\mathbb{P}$, ce qui s'écrit simplement

$$\pi_1 = \pi_0 \cdot \mathbb{P}$$

ou, encore plus simplement $\pi_1 = \pi_0\mathbb{P}$, comme s'il s'agissait du produit de deux nombres (mais ici il s'agit du produit d'un vecteur et d'une matrice). Cette formule très courte signifie que le vecteur $\pi_1 = (\pi_1(h), \pi_1(a), \pi_1(f))$ est le produit du vecteur $\pi_0 = (\pi_0(h), \pi_0(a), \pi_0(f))$ par la matrice $\mathbb{P}$, ce qui s'écrit encore de façon matricielle :

$$(\pi_1(h), \pi_1(a), \pi_1(f)) = (\pi_0(h), \pi_0(a), \pi_0(f)) \begin{pmatrix} 0,5 & 0,45 & 0,05 \\ 0,1 & 0,5 & 0,4 \\ 0 & 0,1 & 0,9 \end{pmatrix}.$$

[1] Si $\Omega = B_1 \cup B_2$ et $B_1 \cap B_2 = \emptyset$, on a $\mathbb{P}(A) = \mathbb{P}(A/B_1)\mathbb{P}(B_1) + \mathbb{P}(A/B_2)\mathbb{P}(B_2)$. Et plus génaralement, si $\Omega = B_1 \cup \ldots \cup B_n$ et $B_i \cap B_j = \emptyset$ si $i \neq j$, on a $\mathbb{P}(A) = \mathbb{P}(A/B_1)\mathbb{P}(B_1) + \ldots + \mathbb{P}(A/B_n)\mathbb{P}(B_n)$.

3 Aperçu de la situation générale

L'exemple de la dynamique d'une forêt naturelle en région tempérée que nous avons étudié dans les deux paragraphes précédents est un cas particulier de *chaîne de Markov.* Plus généralement, on modélise au moyen d'une chaîne de Markov l'évolution au cours du temps de quantités X qui peuvent prendre un nombre fini d'états $X = x_1$, $X = x_2$, ..., $X = x_n$ et qui passent 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 tous les inégalités $0 \leq p_{ij} \leq 1$ et leur somme vaut 1, $\sum_{j=0}^{n} p_{ij} = 1$ (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 $p_{i1} + p_{i2} + \ldots + p_{in} = 1$). L'expression $P(X_{t+1} = x_j / X_t = x_i)$ s'appelle une *probabilité conditionnelle* et représente la "probabilité que la quantité X vaille x_j à l'instant $t+1$ *sachant qu*'elle valait x_i à l'instant t".

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

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

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

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

On appelle *distribution initiale* ou *loi de probabilité initiale* le vecteur π_0

S	x_1	x_2	....	x_n
π_0	$\pi_0(x_1) = P(X_0 = x_1)$	$\pi_0(x_2) = P(X_0 = x_2)$	....	$\pi_0(x_n) = P(X_0 = x_n)$

dont chaque composante représente la probabilités de se trouver dans l'un des états à l'instant initial. La somme des composantes de ce vecteur de probabilités est égale à 1 puisqu'on se trouver de façon certaine dans l'un des états possibles. La connaissance de la matrice de passage $\mathbb{P}$ permet de calculer l'évolution au cours du temps de cette distribution initiale. Pour cela, on fait simplement le produit du vecteur π_0 par cette matrice, une fois pour obtenir la distribution π_1 à l'instant $t = 1$, puis une nouvelle fois pour la connaitre à l'instant $t = 2$ etc...

4 Pertinence des modèles Markoviens

Ajoutons pour finir quelques commentaires concernant la pertinence de la modélisation par des chaines de Markov. Les chaînes de Markov fournissent en effet des modèles très utiles, notamment en biologie mais ils présentent aussi, bien entendu, des défauts. Parmi eux, l'hypothèse simplificatrice d'un **nombre fini d'états** possibles est relativement facile à contourner (lorsque c'est nécessaire) car il existe des chaînes de Markov ayant un nombre infini d'états. De même on peut s'affranchir de l'hypothèse de l'**invariance dans le temps des probabilités de transitions** (on a en effet supposé que la probabilité de passer d'un état x_i à un état x_j ne changeait pas au cours du temps) en considérant des chaînes de Markov qui ne sont pas homogènes, c'est-à-dire ayant des matrices de transition modifiables au cours du temps. Bien entendu l'étude de ces modèles généralisés (chaînes de Markov infinies ou non homogènes) nécessitent le recours à des outils mathématiques plus élaborés.

Par contre, il y a un autre défaut qui, lui est une réelle limite du modèle : c'est l'**invariance spatiale**. Pour le calcul des probabilités de transition, on a fait en effet implicitement une hypothèse d'homogénéité spatiale qui est bien rarement satisfaite dans la pratique. Par exemple un site végétal n'a certainement pas la même probabilité d'être occupé la période suivante par une espèce donnée selon que les sites voisins le sont déjà ou qu'ils ne le sont pas. Et, malheureusement ceci ne peut pas être pris en compte par les modèles markoviens que nous avons présentés ici. On retiendra donc qu'il convient d'utiliser avec prudence une chaîne de Markov lorsque ce caractère d'isotropicité du milieu n'est pas du tout satisfait.

[2]On a coutume de noter les coefficients d'une matrice (p_{ij}), i étant l'indice de ligne et j l'indice de colonne, donc le nombre noté p_{ij} figure dans la i-ème **Li**gne et la j-ème **Co**lonne ; retenez "**LiCo**".

5 Lectures

Pour en savoir plus, il est vivement recommandé de lire l'article de Wikipedia sur les chaines de Markov à l'adresse :

http ://fr.wikipedia.org/wiki/chaine_de_Markov

(et notamment la savoureuse histoire de Doudou le hamster !).