EXEMPLE

Une chaîne de Markov est un cas particulier d'un système dynamique ou on considère plusieurs états d'une seule population. La distinction essentielle est que la population totale est constante. Comme exemple on peut considérer une modèle d'une maladie contagieuse. Le modèle le plus simple considère deux cas seulement: malade et bonne santé.

Exemple 3.1. Chaque mois une personne en bonne santé a une chance de 3% de tomber malade, et une personne atteinte a une chance de 5% de se guérir. Au départ, personne est malade.

On a une matrice de transition (ou matrice d'étape) de $A = \begin{bmatrix} 0,97 & 0,05 \\ 0,03 & 0,95 \end{bmatrix}$. Les colonnes correspondent au deux états au présent et les rangées correspondent au deux états à l'étape suivante.

Puisque la population totale (malade et non-malades) est constante, on considère souvent le vecteur des proportions au lieu du vecteur des populations. Donc on a $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ et $\mathbf{x}_{k+1} = A\mathbf{x}_k$ comme avant. On dit souvent vecteur d'état pour $\mathbf{x}_k$. □

Note que c'est un modèle à 2 états (malade et non-malade), qui mène à une matrice de taille 2×2 et des vecteurs d'état de taille 2. En général, ayant n états on a une matrice de taille $n \times n$ et des vecteurs d'état de taille n.

MATRICE STOCHASTIQUE

On observe que la somme de chaque colonne de la matrice A ci-haut est 1. Ce n'est pas par hasard, et on a même utilisé cette observation afin de trouver les composants non-diagonaux de A. On pourrait même l'adpoter comme définition.

Définition 3.2. Une matrice A est STOCHASTIQUE si les composants de A sont tous non-négatifs et que la somme de chaque colonne est 1. Autrement dit on exige que $A_{ij} \geq 0$ et que $\sum_i A_{ij} = 1$. □

La matrice de transition d'une chaîne de Markov est stochastique (et toute matrice stochastique représente une chaîne de Markov).

On se rappelle que les valeurs propres d'une matrice sont les racines de $\det(A - \lambda I)$. C'est utile de considérer deux théorèmes.

Théorème 3.3. *Les valeurs propres de A sont exactement les mêmes que les valeurs propres de A^T (mais les vecteurs propres sont typiquement différents).* □

Théorème 3.4. *Si la somme de chaque rangée d'une matrice est égale à r (donc, constante) alors le vecteur $\mathbf{1} = (1, 1, \ldots, 1)$ est vecteur propre et $\lambda = r$ est valeur propre.* □

Vous êtes invités à démontrer ces deux théorèmes (on verra un peu en classe).

On applique. Soit A une matrice stochastique. Alors les rangées de A^T ont chacune une somme de 1 (pourquoi?). Donc le théorème nous assure que $\mathbf{1}$ est vecteur propre de A^T et que 1 est valeur propre de A^T. Donc on sait que 1 est valeur propre de A (mais on ne connaît pas le vecteur propre correspondant de A).

Résultat: toute matrice stochastique a 1 comme valeur propre. Donc il existe un vecteur $\mathbf{q}$ tel que $A\mathbf{q} = \mathbf{q}$. On peut choisir $\mathbf{q}$ pour que la somme des composants donne 1, donc on a une distribution d'équilibre, souvent dit VECTEUR D'ÉTAT STATIONNAIRE. On a donc le suivant.

Théorème 3.5. *Tout chaîne de Markov possède un vecteur d'état stationnaire.* □

ÉQUILIBRE ET TENDANCE

Un vecteur d'état stationnaire se trouve comme tout autre vecteur propre.

Exemple 3.6. On détermine le vecteur d'état stationnaire de l'exemple 3.1. On cherche les vecteurs propres correspondant à $\lambda = 1$, donc on solutionne $A - I = \underline{0}$.

$$\begin{bmatrix} 0,97 & 0,05 \\ 0,03 & 0,95 \end{bmatrix} - I = \begin{bmatrix} -0,03 & 0,05 \\ 0,03 & -0,05 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & -5/3 \\ 0 & 0 \end{bmatrix}$$

Ceci donne la solution

$$\mathbf{q} = t \begin{bmatrix} 5/3 \\ 1 \end{bmatrix}$$

On choisit la valeur de t pour que $\mathbf{q}$ représente des proportions: c'est-à-dire on choisit t pour avoir la somme des composants de $\mathbf{q}$ égale à 1. Donc $t = 3/8$ ou

$$\mathbf{q} = \begin{bmatrix} 5/8 \\ 3/8 \end{bmatrix} = \begin{bmatrix} 0,625 \\ 0,375 \end{bmatrix}$$

On interprète: la situation d'avoir $62,5\%$ de la population non-malade et $37,5\%$ malade est stable. □

On se demande la tendance éventuelle des vecteurs d'état $\mathbf{x}_k$.

Exercice 3.7. Montrer que ce n'est jamais le cas que $\mathbf{x}_k$ converge vers $\mathbf{0}$. (indice: montrer que la somme des composants de $\mathbf{x}_k$ est égale à 1 pour $k = 1, 2, 3, \ldots$.) □

C'est déjà une différence entre les chaînes de Markov et les systèmes dynamiques en générale: ici l'extinction est impossible.

Exercice 3.8. Montrer que ce n'est jamais le cas que $\mathbf{x}_k$ augmente sans cesse ("converge vers l'infini"). (indice: montrer que la somme des composants de $\mathbf{x}_k$ est égale à 1 pour $k = 1, 2, 3, \ldots$.) □

Par contre ce n'est pas le cas que pour tout vecteur d'état $\mathbf{x}_0$ et toute matrice stochastique A, $\mathbf{x}_k = A^k\mathbf{x}_0$ converge vers un seul vecteur $\mathbf{q}$.

CHAÎNES DE MARKOV ET VALEURS PROPRES?

Souvent une chaîne de Markov possède les deux propriétés suivantes.

- Il n'existe qu'un seul vecteur d'état stationnaire.
- Peu importe le vecteur d'état initial $\mathbf{x}_0$, les vecteurs $\mathbf{x}_k$ convergent vers $\mathbf{q}$.

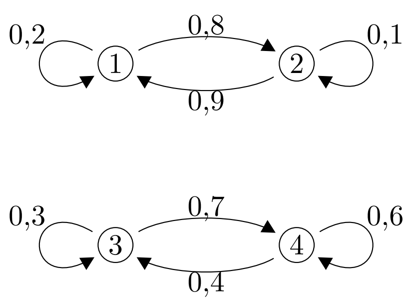
Afin de mieux comprendre ceci, voyons deux exemples où ce n'est pas le cas.

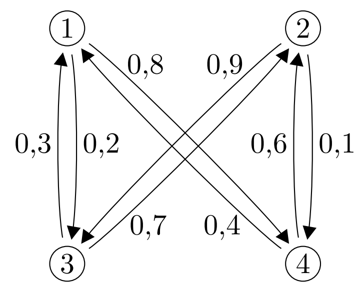
Exercice 3.9. Vérifier que la première matrice de transition possède plus qu'un vecteur d'état stationnaire. Vérifier que pour la deuxième matrice les vecteurs d'état ne convergent pas.

$$\begin{bmatrix} 0,2 & 0,9 & 0 & 0 \\ 0,8 & 0,1 & 0 & 0 \\ 0 & 0 & 0,3 & 0,4 \\ 0 & 0 & 0,7 & 0,6 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 0,3 & 0,4 \\ 0 & 0 & 0,7 & 0,6 \\ 0,2 & 0,9 & 0 & 0 \\ 0,8 & 0,1 & 0 & 0 \end{bmatrix}$$

(indice: les chiffres ne sont pas tellement importants, c'est plutôt les zéros) $\square$

Ce serait utile de considérer les graphes qui correspondent à ces matrices. Pour une chaîne de Markov on construit le graphe en mettant un sommet pour chaque état et une flèche pour chaque transition possible. Voici les graphes correspondants.





Les valeurs propres se montrent utiles ici.

Théorème 3.10. *Soit A une matrice stochastique diagonalisable tel que $\lambda_1 = 1$ est valeur propre de multiplicité 1, et que $|\lambda_j| < 1$ pour toute autre valeur propre λ_j.*

Alors il n'existe qu'un seul vecteur d'état stationnaire $\mathbf{q}$ et peu importe le vecteur d'état initial $\mathbf{x}_0$, les vecteurs $\mathbf{x}_k$ convergent vers $\mathbf{q}$. $\square$

La démonstration de ce théorème est déjà faite: c'est l'équation (2.5) En principe il ne faut que calculer les valeurs propres de A et vérifier que la multiplicité de $\lambda = 1$ est un et que les autre sont tous inférieures à 1 en valeur absolue. En pratique ce n'est si facile: trouver directement les valeurs propres est difficile (parfois impossible?) pour une grande matrice. De plus, il existe des matrices stochastiques qui ne sont pas diagonalisables.

On peut appliquer le théorème 3.10 à l'exercice 3.9. Pour la première matrice, la multiplicité de $\lambda = 1$ est 2 (exercice!). Le vecteur d'état stationnaire n'est pas unique. Pour la deuxième, on voit que $\lambda = 1$ et $\lambda = -1$ sont valeurs propres. Donc l'équation (2.4) est valide mais il faut modifier l'équation (2.5) (voir l'exercice 2.4).

CHAÎNES DE MARKOV RÉGULIÈRES

Il y a une autre approche. On a un théorème qui garantit la convergence désirée, sans devoir calculer les valeurs propres. Une matrice stochastique A est RÉGULIÈRE si il existe un entier ℓ tel que tout composant de A^ℓ est positif.

Théorème 3.11. *Soit A une matrice stochastique régulière. Alors il existe un seul vecteur d'état stationnaire $\mathbf{q}$. De plus pour tout vecteur d'état initial $\mathbf{x}_0$ les vecteurs d'état $\mathbf{x}_k$ convergent vers $\mathbf{q}$.* □

La démonstration est un peu technique. Parfois c'est directement utile.

Exercice 3.12. Montrer que la chaîne de Markov de l'exemple 3.1 converge toujours vers le même vecteur d'état stationnaire. (indice: il faut simplement trouver une valeur de ℓ tel que. . .) □

Parfois il semble utile, mais pas autant qu'on voudrait.

Exemple 3.13. Tenter de calculer le carré de la première matrice A de l'exercice 3.9. On peut montrer, sans calculer tout, que A^2 possède des zéros au même endroits que A (exercice!). On conclut donc qu'il n'existe aucun entier ℓ tel que tout composant de A^ℓ est positif (car A^ℓ possède les mêmes zéros que A).

Par contre le théorème 3.11 ne permet pas de conclure qu'il y a plus qu'un vecteur d'état stationnaire. Ce théorème ne s'applique tout simplement pas, car il s'applique seulement dans les cas ou un ℓ existe. □

Mais ce n'est pas toujours pratique. C'est vrai que $\ell = 1$ ne suffit pas pour les matrices de l'exercice 3.9, mais peut-être que $\ell = 437$ suffit? ou $\ell = 142857$? Le problème c'est qu'on ne connaît pas la valeur de ℓ.

Considérons un graphe (par exemple, le graphe d'une matrice stochastique). Si c'est possible de trouver un chemin de n'importe quel état à n'importe quel autre état, alors on dit que le graphe est FORTEMENT CONNEXE. Un chemin, dans ce sens, doit suivre les flèches dans la "bonne" direction. Un CYCLE est un chemin qui aboutit à son point de départ. Si les longueurs de tous les cycles dans un graphe ont un facteur en commun on dit que le graphes est PÉRIODIQUE (le plus grand facteur commun est la période). Sinon il est apériodique.

Ces définitions mènent au théorème suivant.

Théorème 3.14. *Soit A une matrice stochastique tel que son graphe est fortement connexe et apériodique. Alors A est régulière. Donc, il existe un seul vecteur d'état stationnaire $\mathbf{q}$, et pour tout vecteur d'état initial $\mathbf{x}_0$ les vecteurs d'état $\mathbf{x}_k$ convergent vers $\mathbf{q}$.* □

On verra l'idée de la démonstration en classe. Vous êtes invités de refaire l'exercice 3.9 en utilisant ce théorème!

PUISSANCES

Posons A une matrice stochastique régulière (peut-être qu'on a vérifié que le graphe est fortement connexe et apériodique). Posons $\mathbf{q}$ pour le vecteur d'état stationnaire. Il reste une observation qui est parfois très utile.

On sait que $A^k\mathbf{x}_0 \approx \mathbf{q}$ si k est "grand", peu importe le vecteur initial $\mathbf{x}_0$. Pour faciliter la présentation imaginons un modèle à $n = 3$ états, donc A est de taille 3×3. On peut choisir $\mathbf{x}_0$ comme on veut, donc par exemple:

$$A^k \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \approx \mathbf{q} \qquad A^k \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \approx \mathbf{q} \qquad A^k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \approx \mathbf{q}$$

Exercice 3.15. Que représente les trois produits matriciels? Comme conséquence, que peut-on dire au sujet de A^k? (indice: considérer les colonnes de A^k) □

On (re)découvre dans cette idée une quantification de "grand". Pour une matrice stochastique (régulière), un "grand" k signifie que après k étapes, on est très "proche" du vecteur d'état stationnaire.

Exercice 3.16. Pour une certaine matrice stochastique régulière A, on se demande si après $k = 100$ étapes, ls vecteur d'état $\mathbf{x}_{100}$ est toujours (peu importe $\mathbf{x}_0$) très proche du vecteur d'état stationnaire. On ne vous donne pas A, mais on donne

$$A^{100} = \begin{bmatrix} 0,3769 & 0,3771 & 0,0728 \\ 0,5016 & 0,5019 & 0,0967 \\ 0,1214 & 0,1209 & 0,8304 \end{bmatrix}$$

Expliquer pourquoi $k = 100$ n'est pas assez "grand" pour cette matrice A. (indice: c'est relié à l'exercice 3.15). □

ATTENTION

Un avertissement s'impose. On n'a pas du tout complété l'analyse de la théorie des chaînes de Markov. En particulier, c'est possible d'avoir une chaîne de Markov qui converge toujours vers un vecteur d'état stationnaire, mais n'est pas régulière.

Exercice 3.17. Montrer que la matrice $A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ n'est pas régulière mais on a encore que pour tout vecteur d'état initial $\mathbf{x}_0$, les vecteurs $\mathbf{x}_k = A\mathbf{x}_0$ convergent vers un vecteur d'état stationnaire unique $\mathbf{q}$. Déterminer aussi $\mathbf{q}$.

Pouvez-vous donner d'autres exemples similaires? □