

NOM :
PRENOM :

Date :
Groupe :

Mathématiques Appliquées à la Biologie : Feuille-réponses du TD 3
Initiation au calcul matriciel

On répondra aux questions posées aussi clairement que possible dans les espaces prévus et on remettra cette *feuille-réponses* en fin de séance à l'enseignant chargé du Cours/TD.

Exercice 1. :

1. Soit une chaîne de Markov à 2 états de matrice de transition \mathbb{P} telle que $p_{11} = \frac{3}{4}$ et $p_{22} = \frac{1}{2}$. Calculer l'image π_1 par la chaîne de Markov de la distribution initiale $\pi_0 = (0 ; 1)$.

2. Vérifier que $\pi_0^* = (\frac{2}{3} ; \frac{1}{3})$ est une distribution stationnaire pour cette chaîne de Markov.

3. En déduire que $\lambda = 1$ est une valeur propre à gauche de la matrice \mathbb{P} ; expliquez.

4. Calculer le carré \mathbb{P}^2 . En déduire l'image π_2 de la distribution initiale $\pi_0 = (0 ; 1)$ après deux étapes.

Exercice 2. : Un modèle simpliste¹ d'un brin d'ADN est celui d'un enchaînement "désordonné" de nucléotides de l'un des 4 types adenine (a), cytosine (c), guanine (g) et thymine (t). Le brin est considéré alors comme une trajectoire d'une chaîne de Markov X_n à quatre états $S = \{a, c, g, t\}$ dont la matrice de transition \mathbb{P} fournit les probabilités que l'un de ces états succède à un autre. Ainsi le brin *aagc* est la trajectoire $X_0 = a, X_1 = a, X_2 = g, X_3 = c$.

1. En supposant que la matrice \mathbb{P} est donnée par $\mathbb{P} = \begin{pmatrix} 0 & 0,3 & 0 & 0,7 \\ 0,6 & 0 & 0,4 & 0 \\ 0 & 0,8 & 0 & 0,2 \\ 0,3 & 0 & 0,7 & 0 \end{pmatrix}$ calculer la matrice \mathbb{P}^2 et en déduire la probabilité de trouver dans ce gène une succession de 3 nucléotides *cxt* où $x \in \{a, c, g, t\}$ est l'un quelconque des 4 nucléotides.

2. La distribution initiale $\pi_0 = (\frac{1}{2}; 0; \frac{1}{2}; 0)$ est-elle une distribution stationnaire pour cette chaîne de Markov ?

3. Compléter les valeurs manquantes et en déduire la distribution π_3 si $\pi_0 = (\frac{1}{4}; \frac{1}{4}; \frac{1}{4}; \frac{1}{4})$.

$$\mathbb{P}^3 = \begin{pmatrix} 0 & \dots & 0 & 0,395 \\ 0,45 & \dots & \dots & 0 \\ 0 & 0,53 & 0 & \dots \\ \dots & 0 & 0,505 & 0 \end{pmatrix}$$

4. Supposons que vous puissiez observer un très long brin d'ADN. Indiquer quelle méthode vous choisiriez pour estimer les probabilités de transition figurant dans la matrice \mathbb{P} .

¹Bien que ce modèle soit très peu réaliste il est cependant à la base de nombreuses techniques utilisées en génétique pour reconnaître par exemple les parties codantes d'un gène (exons) des parties non (introns), ces dernières présentant des successions de bases bien plus désordonnées, ou pour comparer une séquence de gène dont on ne connaît pas encore la fonction avec une séquence déjà étudiée.