LIÈGE université
Sciences Appliquées

Faculté des Sciences appliquées

MATH0488-1 Éléments de Processus Stochastiques

Méthodes de Monte Carlo par chaînes de Markov pour la détection de communautés dans un graphe

---

Professeur :
Pierre GEURTS

Groupe :
Maxime FIRRINCIELI
Romain LAMBERMONT
Arthur LOUIS

4 avril 2022

Table des matières

Table des figures

Introduction

Ce projet a pour but de détecter les communautés dans un graphe grâce aux chaînes de Markov et par les méthodes de Monte-Carlo. L'utilisation de ces outils nous permettra de détecter efficacement les communautés dans un graphe. Le projet sera constitué en deux parties. Une première se focalisera sur la familiarisation avec les chaînes de Markov et les méthodes de Monte-Carlo, plus précisement l'algorithme de Metropolis-Hastings, tandis que la seconde sera le cœur du projet avec la recherche de communautés dans un graphe.

1 Première partie : chaînes de Markov et algorithme MCMC

1.1 Chaînes de Markov

Rappelons d'abord brievement ce qu'est une chaîne de Markov. Une chaîne de Markov est un outil mathématique stochastique, qui utilise un principe de "non-mémoire". Tout état d'un système est simplement calculé à partir du précédent, ce qui en facilite l'analyse.

Ces chaînes sont simplement décrites mathématiquement comme suit avec $X_1, X_2, ..., X_t$ une suite de variables aléatoires qui définit une chaîne de Markov si (pour t > 1) elle suit cette relation :

$$\mathbb{P}(X_1, X_2, ..., X_t) = \mathbb{P}(X_1)\prod_{l=2}^{t}\mathbb{P}(X_l|\ X_{l-1})$$

1.1.1

Nous calculons donc pour des valeurs de t croissantes les différentes valeurs demandées, ici avec $t = 20$ (suffit pour avoir convergé) :

- Cas de base distribué uniformément : $\mathbb{P}(X_t = x) = \begin{pmatrix}0.3488 & 0.0698 & 0.2326 & 0.3488\end{pmatrix}$ avec $x = 1, 2, 3, 4$
- Cas de base fixé : $\mathbb{P}(X_t = x) = \begin{pmatrix}0.3488 & 0.0698 & 0.2326 & 0.3488\end{pmatrix}$ avec $x = 1, 2, 3, 4$
- $Q^t = \begin{pmatrix}0.3488 & 0.0698 & 0.2325 & 0.3488\\ 0.3488 & 0.0698 & 0.2326 & 0.3488\\ 0.3488 & 0.0698 & 0.2326 & 0.3488\\ 0.3488 & 0.0698 & 0.2326 & 0.3488\end{pmatrix}$

On remarque donc bien une convergence vers des probabilités et ceci peut importe le cas de départ, on peut montrer cette convergence sur les figures ci-dessous :

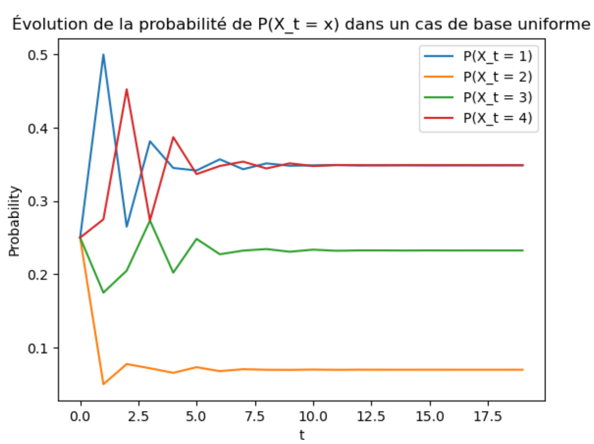

FIGURE 1 – Évolution des probabilités dans une distribution de départ uniforme

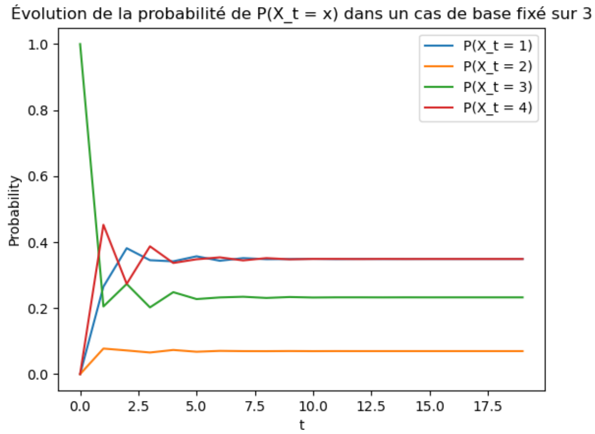

FIGURE 2 – Évolution des probabilités dans une distribution de départ fixée sur 3

1.1.2

Afin de déduire la distribution stationnaire π_∞ de notre chaîne qui est décrite comme suit :

$$[\pi_\infty]_j = \lim_{t\to\infty} \mathbb{P}(X_t = j)$$

Nous allons simplement calculer $\mathbb{P}(X_t)$ avec un grand t ce qui nous donne :

$$\pi_\infty = \begin{pmatrix} 0.3488 & 0.0698 & 0.2326 & 0.3488 \end{pmatrix}$$

1.2 Méthode MCMC : analyse théorique dans le cas fini

1.3 Méthode MCMC : illustration sur un exemple simple

2 Deuxième partie : détection de communautés dans un graphe par algorithmes MCMC

2.1 Etude théorique

2.2 Analyse expérimentale

2.3 Application à un grand graphe