

# Examen M2, cours MAIN

Michel Habib

28 mars 2018

NB : Les exercices sont indépendants, \* signifie une question plus difficile. Tous les documents sauf électroniques sont autorisés.

## 1 Perron-Frobenius

1. Rappeller l'énoncé du théorème de Perron-Frobenius, sur la convergence de PageRank.
2. Montrer sur des exemples que les conditions sont bien nécessaires.
3. \* Montrer que les conditions sont suffisantes.

## 2 Variations sur PageRank

1. Donner au moins 3 familles de graphes sur lesquels PageRank est uniforme (i.e., tous les sommets ont exactement le même PageRank).
2. Proposer une heuristique pour diminuer les écarts de PageRank sur un graphe. On a le droit d'ajouter des arcs préciser lesquels.
3. Si dans l'algorithme de pageRank avec un facteur zap, on choisit  $d=0.8$ , que va-t-il se passer ? En particulier, est-ce que le calcul s'arrête (i.e., convergence normale) ?
4. \* Comment interpréter le résultat obtenu ? Précisez votre intuition.

## 3 Toujours PageRank

On considère un graphe orienté  $G$  qui vérifie les conditions du théorème de Perron-Frobenius, et deux sommets  $x, y \in G$  ayant des voisinages distincts et les mêmes degrés.  $d^+(x) = d^+(y)$  et  $d^-(x) = d^-(y)$ .

1. Soit  $G_{xy}$  le graphe obtenu à partir de  $G$  en contractant les sommets  $x$  et  $y$  en un seul sommet  $z$  et en fusionnant les voisinages.  
 $N^+(z) = N^+(x) \cup N^+(y)$  et  $N^-(z) = N^-(x) \cup N^-(y)$   
Est-ce toujours vrai que  $G_{xy}$  vérifie les conditions du théorème de Perron-Frobenius, si  $G$  les vérifie ?
2. On supposera maintenant que  $G_{xy}$  vérifie toujours les conditions du théorème de Perron-Frobenius et qu'en partant du vecteur uniforme  $1/n$  à chaque ième étape du calcul de PageRank,

$$\text{PageRank}_i(x) = \text{PageRank}_i(y)$$

Que vaut le PageRank de  $z$  calculé dans  $G_{xy}$  en fonction des PageRanks de  $x$  et  $y$  calculés dans  $G$  ?

(Piste : on peut commencer par considérer le cas où le vecteur initial est  $1/n$ .)

3. La condition : voisinages distincts est-elle nécessaire pour le résultat de la question précédente ?
4. En déduire le PageRank du graphe suivant, constitué de trois circuits de longueurs différentes 3, 5, 7 ayant un sommet en commun.
5. Quelle est la classe de graphes orientés sur laquelle cette opération permet de calculer le PageRank facilement ?

## 4 Spams

1. Précisez les conditions qui caractérisent une page web dite de spam (cf. cours).
2. \* En déduire un algorithme de classification des pages Web. On pourra fournir un indice de spam entre 0 et 1.
3. Comment définir un mail qui est un spam ?
4. Comment s'en servir pour améliorer la méthode de la question 2 de l'exercice.

## 5 Analyse à la volée de Twitter

1. Dans le cadre de la surveillance généralisée on vous demande de faire une application de contrôle de Twitter. Qu'allez-vous contrôler et avec quels algorithmes ?
2. Comment reconnaître une FakeNews dans twitter ?