Modèles et algorithmes des réseaux

15 septembre

Exercice 1 - PageRank et le temps de mélange

On définit la distance en variation totale entre deux mesures de probabilité μ et ν sur Ω (fini) par :

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subset \Omega} |\mu(A) - \nu(A)|.$$

1. Montrer que
$$\frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \sum_{\mu(x) \geq \nu(x)} (\mu(x) - \nu(x)).$$

2. Montrer que pour $B = \{x, \mu(x) \geq \nu(x)\}$ et tout $A \subset \Omega$, on a :
$$\mu(A) - \nu(A) \leq \mu(B) - \nu(B).$$

3. En déduire que
$$\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Un couplage entre deux mesures de probabilité μ et ν est une paire de variables aléatoires (X, Y) définies sur un même espace de probabilité telle que la distribution marginale de X est μ, i.e. $\mathbb{P}(X = x) = \mu(x)$ et la marginale de Y est ν, i.e. $\mathbb{P}(Y = y) = \nu(y)$.

4. Montrer que
$$\|\mu - \nu\|_{\mathrm{TV}} \leq \inf\{\mathbb{P}(X \neq Y) \ : \ (X, Y) \text{ est un couplage de } \mu \text{ et } \nu\}.$$

5. Montrer que
$$\sum_{x \in \Omega} (\mu(x) \wedge \nu(x)) = 1 - \|\mu - \nu\|_{\mathrm{TV}} \in [0, 1].$$

6. En déduire un couplage (X, Y) tel que $\|\mu - \nu\|_{\mathrm{TV}} = \mathbb{P}(X \neq Y)$.

Une chaîne de Markov (homogène) à espace d'états Ω et de matrice de transition P est une suite de variables aléatoires $(X_0, X_1, \dots)$ telles que

$$\mathbb{P}(X_{t+1} = y | X_t = x, X_{t-1} = x_{t-1}, \dots, X_0 = x_0) = \mathbb{P}(X_{t+1} = y | X_t = x) = P(x, y).$$

Une chaîne est dite irréductible si pour tous $x, y \in \Omega$, il existe t tel que $P^t(x, y) > 0$. Soit $T(x) = \{t \geq 1, P^t(x, x) > 0\}$. La période de x est le pgcd de $T(x)$. Une chaîne est dite apériodique si tous les états ont pour période 1.

7. Montrer que si P est irréductible alors tous les états ont même période.

8. Montrer que si P est irréductible apériodique alors P est primitive.
(Note : tout ensemble des entiers non-négatifs fermé sous addition et avec un pgcd égal à 1 contient tous sauf un nombre fini d'entiers.)
9. Montrer que dans ce cas, il existe une unique distribution stationnaire sur Ω, satisfaisant $\pi = \pi P$.

On note $P^t(x, .)$ la loi de la chaîne de Markov au temps t commencé en $X_0 = x \in \Omega$ et μP^t pour le cas où la loi de X_0 est μ. On définit

$$d(t) = \max_{x \in \Omega} \|P^t(x, .) - \pi\|_{\mathrm{TV}},$$
$$\overline{d}(t) = \max_{x,y \in \Omega} \|P^t(x, .) - P^t(y, .)\|_{\mathrm{TV}}.$$

10. Montrer que
$$d(t) \leq \overline{d}(t) \leq 2d(t).$$
11. Montrer que
$$d(t) = \sup_{\mu} \|\mu P^t - \pi\|_{\mathrm{TV}},$$
$$\overline{d}(t) = \sup_{\mu,\nu} \|\mu P^t - \nu P^t\|_{\mathrm{TV}},$$
où μ, ν sont des distributions sur Ω.
12. Montrer que
$$\|\mu P - \nu P\|_{\mathrm{TV}} \leq \|\mu - \nu\|_{\mathrm{TV}}.$$
En déduire que $d(t)$ et $\overline{d}(t)$ sont décroissantes en t.
13. Pour un couplage (X_s, Y_s) de $P^s(x, .)$ et $P^s(y, .)$, montrer que
$$\|P^{t+s}(x, .) - P^{t+s}(y, .)\|_{\mathrm{TV}} = \frac{1}{2} \sum_z |\mathbb{E}[P^t(X_s, z) - P^t(Y_s, z)]|.$$
14. En déduire que $\overline{d}(t)$ est sous-multiplicative :
$$\overline{d}(t)(t+s) \leq \overline{d}(t)\overline{d}(s).$$

On définit le temps de mélange par
$$t_{\mathrm{mix}}(\epsilon) = \min\{t, d(t) \leq \epsilon\}$$
et
$$t_{\mathrm{mix}} = t_{\mathrm{mix}}(1/4).$$

15. Montrer que $d(kt_{\mathrm{mix}}) \leq 2^{-k}$ et $t_{\mathrm{mix}}(\epsilon) \leq \lceil \log_2 \epsilon^{-1} \rceil t_{\mathrm{mix}}$.

On définit un couplage de chaînes de Markov avec matrice de transition P comme le processus $(X_t, Y_t)_{t \geq 0}$ ayant la propriété que chacun des processus (X_t) et (Y_t) est une chaîne de Markov de matrice de transition P (ces processus peuvent avoir des points de départs différents). Tout couplage peut être modifié de telle sorte que : si $X_s = Y_s$ alors $X_t = Y_t$ pour tout $t \geq s$. On définit alors le temps de couplage par

$$\tau_{\mathrm{couple}} = \min\{t, X_t = Y_t\}.$$

16. Comme exemple, considérer la marche aléatoire sur $\{0, 1, \ldots, n\}$ qui monte ou descend avec probabilité 1/2 et qui, aux bords, reste immobile avec probabilité 1/2. En construisant un couplage, montrer que $P^t(x, n) \leq P^t(y, n)$ dès que $x \leq y$.
17. Montrer que
$$\|P^t(x, .) - P^t(y, .)\|_{\mathrm{TV}} \leq \mathbb{P}_{x,y}(\tau_{\mathrm{couple}} > t).$$
18. Considérer la chaîne de Markov associée au calcul de Page Rank (qui choisit avec probabilité $1 - \theta$ un sommet uniformément au hasard à chaque étape). Montrer que son temps de mélange satisfait $t_{\mathrm{mix}} \leq \frac{4}{1-\theta}$.
Qu'en déduire sur la vitesse de convergence de l'algorithme ?

Exercice 2 - PageRank : algèbre linéaire

Comme dans le cours, on note M la matrice carrée telle que $M_{ij} = \frac{1}{d^+(i)}$ si $i \to j$ et zero sinon. Si i est une feuille, i.e. n'a pas de lien sortant alors la ligne i de la matrice M est nulle. On définit $\overline{M}$ la matrice carrée où chacune de ces lignes nulles est remplacée par une ligne $(\frac{1}{n}, \dots, \frac{1}{n})$.

1. Montrer que 1 est valeur propre de $\overline{M}$ et toute valeur propre (complexe) λ de $\overline{M}$ vérifie $|\lambda| \leq 1$.

On définit alors $\tilde{M} = \alpha\overline{M} + (1-\alpha)\mathbf{1}\mathbf{p}^T$ où $\mathbf{p}$ est un vecteur de probabilité. Remarquer que le cas $\mathbf{p} = \frac{1}{n}\mathbf{1}$ est le cas vu en cours.

2. Montrer qu'il existe une matrice non-singulière $P = (\mathbf{1}\ X)$ telle que :

$$P^{-1}\overline{M}P = \begin{pmatrix} 1 & A \\ 0 & B \end{pmatrix}.$$

3. En utilisant P, montrer que les valeurs propres de $\tilde{M}$ sont $\{1, \alpha\lambda_2, \dots, \alpha\lambda_n\}$ où $\{1, \lambda_2, \dots, \lambda_n\}$ sont les valeurs propres de $\overline{M}$.
4. Conclure sur la vitesse de convergence de l'algorithme PageRank.

On définit le vecteur $\mathbf{a}$ par $a_i = 1$ si la ligne i de M correspond à une feuille, et 0 sinon. Montrer que l'algorithme PageRank calcule π qui satisfait $\pi^T(I - \alpha M) = \mathbf{p}^T$ avec $\pi^T\mathbf{1} = 1$.

5. Montrer que $I - \alpha M$ est inversible pour $\alpha < 1$.
6. En écrivant

$$M = \begin{pmatrix} M_{11} & M_{12} \\ 0 & 0 \end{pmatrix},$$

montrer que $I - \alpha M_{11}$ est inversible puisque l'algorithme suivant calcule le vecteur π :

(a) résoudre : $\pi_1^T(I - \alpha M_{11}) = \mathbf{p_1}^T$.

(b) calculer : $\pi_2^T = \alpha\pi_1^T P_{12} + v_2^T$.

(c) normaliser $\pi = \frac{1}{\|\pi_1\|_1 + \|\pi_2\|_1}(\pi_1, \pi_2)$.

Quel est l'intérêt de cet algorithme ?