Coloriages aléatoires

Université de Grenoble - Préparation à l'agrégation

Christophe Leuridan

Lorsqu'on colorie une carte du monde, on souhaite que deux pays ayant une frontière commune soient de couleur différentes. Le théorème des quatre couleurs affirme que quatre couleurs suffisent à colorier toute carte plane pourvu que les pays soient d'un seul tenant.

Ce problème est un cas particulier du coloriage d'un graphe fini : il suffit d'associer un sommet à chaque pays et de relier deux sommets par une arête si les pays correspondant ont une frontière commune. Dans ce texte, nous étudions un algorithme pour simuler le coloriage aléatoire d'un graphe.

1 Introduction : le polynôme chromatique

Un graphe fini non orienté $\mathcal{G}$ est la donnée d'un ensemble fini $\mathcal{V}$ et d'une partie $\mathcal{E}$ de l'ensemble $\mathcal{P}_2(\mathcal{V})$ des parties de $\mathcal{V}$ à deux éléments. Les éléments de $\mathcal{V}$ sont les sommets (en anglais « vertices ») et les éléments de $\mathcal{E}$ sont les arêtes (en anglais « edges »).

Si u et v sont deux sommets distincts, on dira que u et v sont voisins dans $\mathcal{G}$ et on notera $u \sim_G v$ lorsque $\{u, v\} \in \mathcal{E}$. On appellera degré d'un sommet v dans le graphe $\mathcal{G}$, noté d_v, le nombre de voisins de v dans $\mathcal{G}$. On notera Δ le degré maximum.

On se donne un ensemble fini $\mathcal{C}$ de couleurs. Un coloriage admissible de $\mathcal{G}$ par $\mathcal{C}$ est une application x de $\mathcal{V}$ dans $\mathcal{C}$ telle que pour tout $u, v \in \mathcal{V}$, $u \sim_G v$ entraîne $x(u) \neq x(v)$. On notera $A(\mathcal{G}, \mathcal{C})$ l'ensemble des coloriages admissibles de $\mathcal{G}$ par $\mathcal{C}$. Dénombrer le nombre d'éléments de $A(\mathcal{G}, \mathcal{C})$ est souvent difficile, en dehors de quelques graphes particulièrement simples. On dispose toutefois d'un résultat théorique remarquable.

Théorème 1 *Soit $\mathcal{G}$ un graphe fini non orienté. Il existe un unique polynôme $P_\mathcal{G}$ appelé polynôme chromatique tel que pour tout ensemble fini $\mathcal{C}$ de couleurs,*

$$\#A(\mathcal{G}, \mathcal{C}) = \mathcal{P}_\mathcal{G}(\#\mathcal{C}).$$

De plus, $P_\mathcal{G}(q) > 0$ dès que $q \geq \Delta + 1$ et $P_\mathcal{G}(q) \sim q^{\#\mathcal{V}}$ quand $q \to +\infty$.

L'existence du polynôme chromatique se montre par récurrence sur le nombre d'arêtes du graphe $\mathcal{G}$. Etant donné une arête e d'un graphe $\mathcal{G}$ reliant les sommets v_1 et v_2, on peut considérer le graphe $\mathcal{G} \setminus e$ obtenu en supprimant l'arête e et le graphe $\mathcal{G}/e$ obtenu en contractant l'arête e : les deux sommets v_1 et v_2 sont remplacés par un seul sommet v dans le graphe $\mathcal{G}$, et les voisins de v dans $\mathcal{G}/e$ sont les voisins de v_1 ou de v_2 dans $\mathcal{G}$ autres que v_1 et v_2. Les graphes $\mathcal{G} \setminus e$ et $\mathcal{G}/e$ possèdent moins d'arêtes que $\mathcal{G}$ et on vérifie facilement que $P_\mathcal{G}(q) = P_{\mathcal{G}\setminus e}(q) - P_{\mathcal{G}/e}(q)$.

Suggestions de développement :

1. Compléter la démonstration du théorème 1.
2. Montrer que $P_{\mathcal{G}}(\Delta)$ peut être nul.
3. Calculer $P_G(n)$ pour le graphe complet à N sommets, défini par $\mathcal{V} = [1 \dots N]$ et $\mathcal{E} = \mathcal{P}_2(\mathcal{V})$. Comment simuler un coloriage aléatoire de ce graphe ?
4. Calculer $P_{\mathcal{G}}(n)$ pour le graphe défini par $\mathcal{V} = [1 \dots N]$ et $\mathcal{E} = \{\{k, k+1\}; k \in [1 \dots N-1]\}$. Comment simuler un coloriage aléatoire de ce graphe ?
5. Calculer $P_{\mathcal{G}}(n)$ pour le graphe défini par $\mathcal{V} = [1 \dots N]$ et $\mathcal{E} = \{\{k, k+1\}; k \in [1 \dots N-1]\} \cup \{1, n\}$.

2 Une chaîne de Markov sur les coloriages

On fixe désormais le graphe $\mathcal{G}$ et l'ensemble $\mathcal{C}$ des couleurs. Comme l'ensemble $A = A(\mathcal{G}, \mathcal{C})$ de tous les coloriages admissibles est souvent difficile à décrire et à dénombrer, il est difficile de simuler directement la loi uniforme sur A. Une idée est de simuler une chaîne de Markov sur l'ensemble $\mathcal{C}^{\mathcal{V}}$ des coloriages (admissibles ou non) dont la mesure invariante est la loi uniforme sur A.

On construit par récurrence une chaîne de Markov $(X_n)_{n \in \mathbf{N}}$ de la façon suivante. On part d'un coloriage $X_0 \in \mathcal{C}^{\mathcal{V}}$. Pour passer du coloriage X_n au coloriage X_{n+1},
- on choisit au hasard un sommet V_{n+1} ;
- on choisit au hasard une couleur C_{n+1} ;
- si tous les voisins de V_{n+1} ont une couleur différente de C_{n+1}, on change la couleur de V_{n+1} en C_{n+1}.

Par conséquent, le coloriage X_{n+1} diffère du coloriage X_n en au plus un point, V_{n+1}, et $X_{n+1}(V_{n+1}) = C_{n+1}$ si $X_n(u) \neq C_{n+1}$ pour tout u voisin de V_{n+1}. Sinon, $X_{n+1} = X_n$.

Les probabilités de transition sont simples à décrire : si un coloriage y diffère de x en un seul sommet v et si $y(u) \neq y(v)$ pour tout voisin u de v, alors $p(x, y) = (\#\mathcal{V}\#C)^{-1}$. Les autres probabilités de transition sont nulles à l'exception des probabilités $p(x, x)$. On vérifie facilement que si $A \neq \emptyset$, la loi uniforme sur A est une probabilité invariante de la chaîne de Markov $(X_n)_{n \in \mathbf{N}}$.

L'ensemble des coloriages admissibles possède des propriétés remarquables pour cette chaîne de Markov.

Proposition 2 *(Propriétés de la partie A)*

1. *La partie A est absorbante : si $X_0 \in A$, alors $X_n \in A$ pour tout $n \in \mathbf{N}$.*
2. *Si $\#\mathcal{C} \geq \Delta + 1$, alors la partie A est accessible depuis tout état de $\mathcal{C}^{\mathcal{V}}$.*
3. *Si $\#\mathcal{C} \geq \Delta + 2$, alors tout état de A est accessible depuis tout état de $\mathcal{C}^{\mathcal{V}}$.*

Donnons une idée de la démonstration du point 3, le plus délicat. Pour passer d'un état $x \in A$, à un état $y \in A$ autre que x, on choisit un sommet v_1 où les deux coloriages diffèrent. Notons $N(v_1)$ (comme « neighbourhood ») l'ensemble des voisins de v_1. Si $x(u) \neq y(v_1)$ pour tout $u \in N(v_1)$, on peut directement changer la couleur de v_1 en $y(v_1)$. Sinon, il faut au préalable changer les couleurs des voisins u de v_1 tels que $x(u) = y(v_1)$. Le fait que $\#\mathcal{C} \geq \Delta + 2$ permet de remplacer ces couleurs une à une par une couleur différente de $y(v_1)$. Comme cette opération ne modifie pas les couleurs des sommets où x et y coïncident, elle aboutit à un coloriage plus proche de y.

De ces propriétés, on déduit facilement le résultat suivant.

Théorème 3 *Si $\#\mathcal{C} \geq \Delta + 2$, la loi de X_n tend vers la loi uniforme sur A.*

Suggestions de développement :

1. Pourquoi la suite $(X_n)_{n\in\mathbf{N}}$ est-elle une chaîne de Markov ?
2. Justifier et compléter le calcul des probabilités de transition.
3. Montrer que la loi uniforme sur A est une probabilité invariante. Est-ce la seule ?
4. Justifier les résultats concernant la partie A. Montrer si $\#\mathcal{C} = \Delta + 1$, tous les états de A ne communiquent pas nécessairement entre eux.
5. Que peut-on en déduire sur la nature des états (récurrence, transience) et le comportement asymptotique de la suite $(X_n)_{n\in\mathbf{N}}$?
6. Justifier la convergence en loi. Que sait-on sur la vitesse de convergence ?
7. Simuler la chaîne de Markov sur un graphe simple.

3 Un couplage

Dans ce paragraphe, on se propose d'étudier l'influence des conditions initiales sur la chaîne de Markov. Pour cela on se donne des variables aléatoires indépendantes X_0 et Y_0 à valeurs dans $\mathcal{C}^{\mathcal{V}}$, $V_1, V_2, \ldots$ de loi uniforme sur $\mathcal{V}$ et $C_1, C_2, \ldots$ de loi uniforme sur $\mathcal{C}$ et on définit les chaînes de Markov $(X_n)_{n\in\mathbf{N}}$ et $(Y_n)_{n\in\mathbf{N}}$ par les relations de récurrence $X_{n+1} = f(X_n, V_{n+1}, C_{n+1})$ et $Y_{n+1} = f(Y_n, V_{n+1}, C_{n+1})$, en notant $f(x, v, c)$ le coloriage obtenu à partir de x en changeant la couleur de v en c si les voisins de v ont une couleur différente de c.

On définit la distance de Hamming entre deux coloriages x et y en notant $d(x, y)$ le nombre de sommets où les coloriages x et y diffèrent. Pour tout $n \in \mathbf{N}$, notons $D_n = d(X_n, Y_n)$ et $\mathcal{F}_n = \sigma(X_0, Y_0, \ldots, X_n, Y_n)$. Nous allons étudier le comportement de la suite $(D_n)_{n\in\mathbf{N}}$.

Théorème 4 *Pour tout $n \in \mathbf{N}$,*

$$\mathbf{E}[D_{n+1}|\mathcal{F}_n] \leq \Big(1 + \frac{(3\Delta - \#\mathcal{C})}{\#\mathcal{V}\#\mathcal{C}}\Big) D_n.$$

Par conséquent, si $\#\mathcal{C} \geq 3\Delta$, $(D_n)_{n\in\mathbf{N}}$ est une surmartingale positive, qui converge presque sûrement vers 0.

Démonstration. Démontrons l'inégalité du théorème. Il s'agit de montrer que quels que soient les coloriages x et y,

$$\mathbf{E}\big[d\big(f(x, V_{n+1}, C_{n+1}), f(y, V_{n+1}, C_{n+1})\big)\big] \leq \Big(1 + \frac{(3\Delta - \#\mathcal{C})}{\#\mathcal{V}\#\mathcal{C}}\Big) d(x, y). \qquad (1)$$

Soit $l = d(x, y)$. Choisissons un chemin $x = x_0, \ldots, x_l = y$ de longueur minimale joignant x à y : pour tout $k \in [1 \ldots l]$, les coloriages x_{k-1} et x_k diffèrent en un seul sommet. Pour montrer l'inégalité (1), il suffit de montrer l'inégalité analogue pour chaque couple (x_{k-1}, x_k). Il suffit donc de montrer l'inégalité (1) dans le cas particulier où $d(x, y) = 1$.

Supposons donc que $d(x,y)=1$ et notons v l'unique sommet où les coloriages x et y diffèrent.

La transformation $f(\cdot,V_{n+1},C_{n+1})$ modifie chaque coloriage x et y en au plus un sommet, V_{n+1}, dont la couleur devient alors C_{n+1}. La distance entre les coloriages $f(x,V_{n+1},C_{n+1})$ et $f(y,V_{n+1},C_{n+1})$ ne peut donc valoir que 0, 1 ou 2.

Pour que $d(f(x,V_{n+1},C_{n+1}),f(y,V_{n+1},C_{n+1}))=2$, il faut qu'un seul des coloriages x et y soit modifié. Cela impose que V_{n+1} soit voisin de v et que C_{n+1} soit l'une des couleurs $x(v)$ ou $y(v)$. La probabilité de cet événement est au plus $(2\Delta)/(\#\mathcal{V}\#\mathcal{C})$.

Pour que $d(f(x,V_{n+1},C_{n+1}),f(y,V_{n+1},C_{n+1}))=0$, il suffit que $V_{n+1}=v$ et que C_{n+1} soit différent des couleurs des voisins de v. La probabilité de cet événement est au moins $(\#\mathcal{C}-\Delta)/(\#\mathcal{V}\#\mathcal{C})$.

Par différence, on obtient

$$\mathbf{E}\big[d(f(x,V_{n+1},C_{n+1}),d(f(y,V_{n+1},C_{n+1})\big]-1\leq\frac{(3\Delta-\#\mathcal{C})}{\#\mathcal{V}\#\mathcal{C}}\ ,$$

d'où on déduit l'inégalité annoncée.

Le théorème fournit même une minoration de la vitesse de convergence vers la loi uniforme sur A lorsque $\#\mathcal{C}>3\Delta$. En effet, notons P_Z la loi d'une variable aléatoire Z, π la loi uniforme sur A et lorsque μ et ν sont deux probabilités sur A, notons $d_{VT}(\mu,\nu)$ la distance en variation totale entre μ et ν :

$$d_{VT}(\mu,\nu)=\sup_{B\subset\mathcal{C}^{\mathcal{V}}}|\mu(B)-\nu(B)|=\frac{1}{2}\sum_{x\in\mathcal{C}^{\mathcal{V}}}|\mu\{x\}-\nu\{x\}|.$$

Supposons que Y_0 ait pour loi π. Alors pour tout $n\in\mathbf{N}$, Y_n a aussi pour loi π. En remarquant que

$$\mathbf{I}_{[X_n\neq Y_n]}=\frac{1}{2}\sum_{x\in\mathcal{C}^{\mathcal{V}}}\Big|\mathbf{I}_{[X_n=x]}-\mathbf{I}_{[Y_n=x]}\Big|,$$

on montre facilement que

$$d_{VT}(P_{X_n},\pi)\leq P[X_n\neq Y_n]\leq\mathbf{E}[D_n]\leq\mathbf{E}[D_0]\Big(1-\frac{(3\Delta-\#\mathcal{C})}{\#\mathcal{V}\#\mathcal{C}}\Big)^n\leq\#\mathcal{V}\Big(1-\frac{(3\Delta-\#\mathcal{C})}{\#\mathcal{V}\#\mathcal{C}}\Big)^n.$$

Suggestions de développement :

1. Montrer que la suite $(X_n,Y_n)_{n\in\mathbf{N}}$ est une chaîne de Markov.
2. Compléter la démonstration du théorème
3. Justifier la majoration de $d_{VT}(P_{X_n},\pi)$. Cette majoration vous semble-t-elle optimale ?
4. Soit $D=\{(x,x)\ ;\ x\in\mathcal{C}^{\mathcal{V}}\}$ la diagonale de $\mathcal{C}^{\mathcal{V}}\times\mathcal{C}^{\mathcal{V}}$. Montrer que si $\#\mathcal{C}>\Delta$, la chaîne $(X_n,Y_n)_{n\in\mathbf{N}}$ finit presque sûrement dans D à partir d'un certain rang. En déduire que la loi de X_n tend vers la loi uniforme sur A (si A est non vide).

Références

Ce texte est inspiré de l'article de Alan Frieze et Eric Vigoda « A survey on the use of Markov Chains to randomly sample colouring », pages 53-71 du livre livre édité par Geoffrey Grimmett et Colin McDiarmid « Combinatorics, Complexity, and Chance : A tribute to Dominique Welsh », Oxford lecture series in mathematics and its applications (34), 2007.