Chaînes de Markov, Génération Aléatoire, et Couplage Depuis le Passé (2)

Petite École de Combinatoire

P. Duchon

LaBRI

Résumé de l'épisode précédent

- **Définition des chaînes de Markov (homogènes) :**
  distribution de départ et probabilités de transition;
  marche aléatoire biaisée dans un graphe orienté;

Résumé de l'épisode précédent

- **Définition des chaînes de Markov (homogènes) :** distribution de départ et probabilités de transition; marche aléatoire biaisée dans un graphe orienté;
- **États et ensembles d'états récurrents :** pour un ensemble fini d'états, puits du graphe des composantes fortement connexes; pour un ensemble dénombrable, ce n'est qu'une condition nécessaire et il peut ne pas y avoir d'états récurrents;

Résumé de l'épisode précédent

- **Définition des chaînes de Markov (homogènes) :** distribution de départ et probabilités de transition; marche aléatoire biaisée dans un graphe orienté;
- **États et ensembles d'états récurrents :** pour un ensemble fini d'états, puits du graphe des composantes fortement connexes; pour un ensemble dénombrable, ce n'est qu'une condition nécessaire et il peut ne pas y avoir d'états récurrents;
- **Autre caractérisation des états récurrents :** si la chaîne passe par un état récurrent, elle y passe une infinité de fois (avec probabilité 1); pour un état transient, elle y passe (avec probabilité 1) un nombre fini de fois.

Rappel de notations

- $p_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, soit en v au temps n;

Rappel de notations

- $p_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, soit en v au temps n;
- $q_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, passe pour la première fois en v au temps n;

Rappel de notations

- $p_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, soit en v au temps n;
- $q_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, passe pour la première fois en v au temps n;
- $P_{uv}(t) = \sum_n p_{uv}^{(n)} t^n$, $Q_{uv}(t) = \sum_n q_{uv}^{(n)} t^n$;

Rappel de notations

- $p_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, soit en v au temps n;
- $q_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, passe pour la première fois en v au temps n;
- $P_{uv}(t) = \sum_n p_{uv}^{(n)} t^n$, $Q_{uv}(t) = \sum_n q_{uv}^{(n)} t^n$;
- $P_{uv}(t) = \delta_{uv} + Q_{uv}(t) P_{vv}(t)$;

Rappel de notations

- $p_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, soit en v au temps n;
- $q_{uv}^{(n)}$: probabilité que la chaîne, partant de u au temps 0, passe pour la première fois en v au temps n;
- $P_{uv}(t) = \sum_n p_{uv}^{(n)} t^n$, $Q_{uv}(t) = \sum_n q_{uv}^{(n)} t^n$;
- $P_{uv}(t) = \delta_{uv} + Q_{uv}(t) P_{vv}(t)$;
- u transient si $P_{uu}(1) < +\infty$, récurrent sinon.

États “positifs”

Définition On dira qu’un état u est positif si l’on a

$$\tau_u = \mathbb{E}(T_u | X_0 = u) < +\infty$$

États “positifs”

Définition On dira qu'un état u est positif si l'on a

$$\tau_u = \mathbb{E}(T_u | X_0 = u) < +\infty$$

Remarques :

- T_u est ici le temps de premier retour en u (la chaîne part de u);

États “positifs”

Définition On dira qu’un état u est positif si l’on a

$$\tau_u = \mathbb{E}(T_u | X_0 = u) < +\infty$$

Remarques :

- T_u est ici le temps de premier retour en u (la chaîne part de u);
- Ça implique que u soit récurrent (puisque u est récurrent si et seulement si T_u est presque sûrement fini pour la chaîne partant de u)

États “positifs”

Définition On dira qu’un état u est positif si l’on a

$$\tau_u = \mathbb{E}(T_u | X_0 = u) < +\infty$$

Remarques :

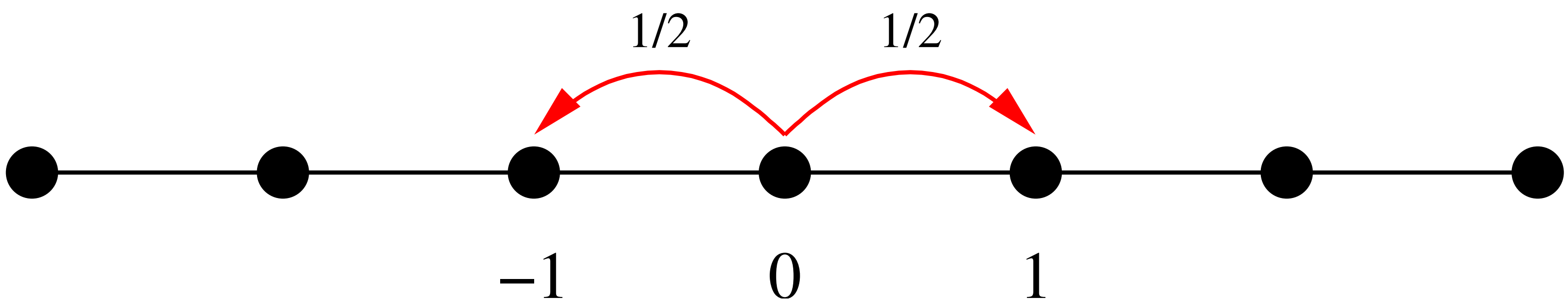
- T_u est ici le temps de premier retour en u (la chaîne part de u);
- Ça implique que u soit récurrent (puisque u est récurrent si et seulement si T_u est presque sûrement fini pour la chaîne partant de u)
- L’espérance s’exprime au moyen de Q'_{uu} :

$$\mathbb{E}(T_u | X_0 = u) = \sum n q_{uu}^{(n)} = Q'_{uu}(1)$$

(récurrence : CV de $P_{uu}(1)$; positivité : CV de $Q'_{uu}(1)$)

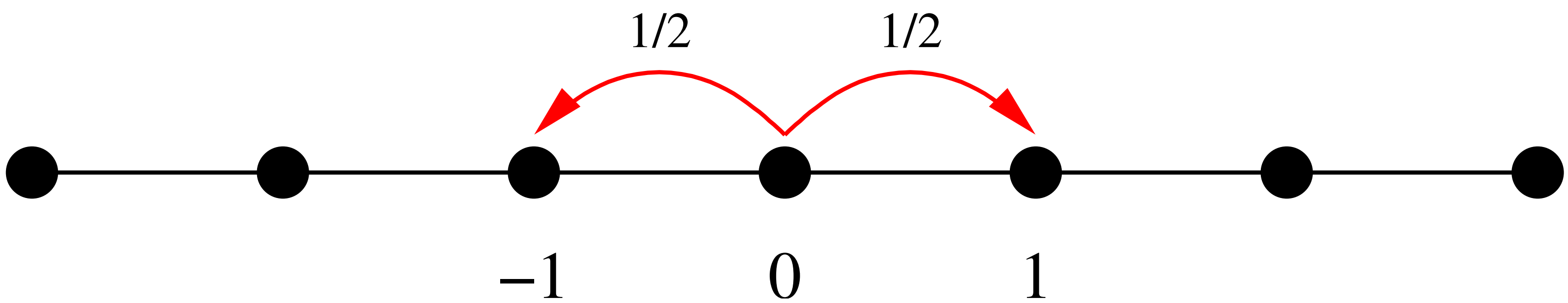
Récurrence et positivité

Pour une chaîne à espace d'états dénombrable, il y a *a priori* une différence entre récurrence et positivité. . .



Récurrence et positivité

Pour une chaîne à espace d'états dénombrable, il y a *a priori* une différence entre récurrence et positivité…

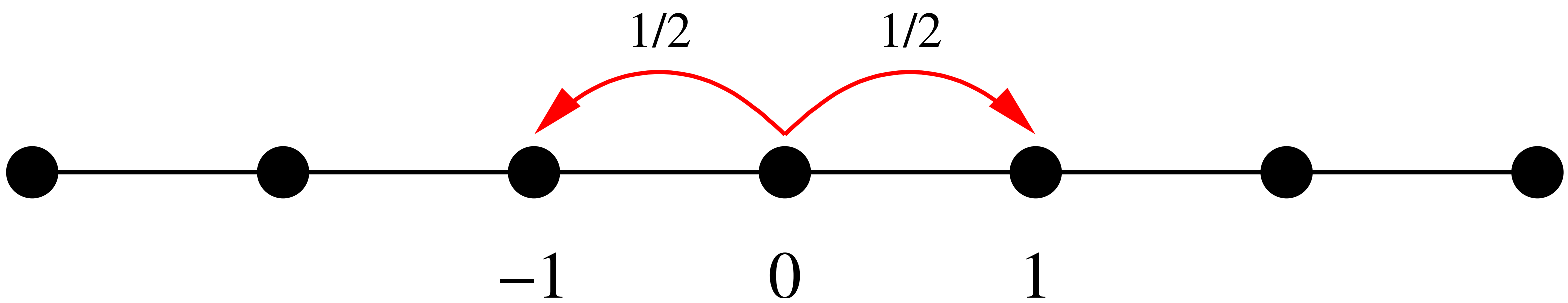


- 0 est récurrent :
  $\mathbb{P}(\text{premier retour en 0 en } 2n+2 \text{ pas}) = 2C_n/2^{2n+2}$, et

$$\sum_{n\geq 0} C_n \left(\frac{1}{4}\right)^n = 2$$

Récurrence et positivité

Pour une chaîne à espace d'états dénombrable, il y a *a priori* une différence entre récurrence et positivité...



- 0 est récurrent :
  $\mathbb{P}($premier retour en 0 en $2n+2$ pas$) = 2C_n/2^{2n+2}$, et

$$\sum_{n\geq 0} C_n \left(\frac{1}{4}\right)^n = 2$$

- 0 n'est pas positif : $\mathbb{E}(T_0) = \sum_n (2n+2).C_n/2^{2n+1} = +\infty$

Critère de positivité

Pour tout état u, on appelle période de u, et on note d_u, le pgcd des n tels que $p_{uu}^{(n)} > 0$.

Critère de positivité

Pour tout état u, on appelle période de u, et on note d_u, le pgcd des n tels que $p_{uu}^{(n)} > 0$. C'est aussi le pgcd des longueurs de cycles dans la composante fortement connexe contenant u.

Critère de positivité

Pour tout état u, on appelle période de u, et on note d_u, le pgcd des n tels que $p_{uu}^{(n)} > 0$. C'est aussi le pgcd des longueurs de cycles dans la composante fortement connexe contenant u. En particulier, si un sommet possède une boucle ($p_{uu} > 0$), tout sommet de sa composante fortement connexe a période 1.

Critère de positivité

Pour tout état u, on appelle période de u, et on note d_u, le pgcd des n tels que $p_{uu}^{(n)} > 0$. C'est aussi le pgcd des longueurs de cycles dans la composante fortement connexe contenant u. En particulier, si un sommet possède une boucle ($p_{uu} > 0$), tout sommet de sa composante fortement connexe a période 1.

Théorème :

- Si u est positif, alors $\lim_n p_{uu}^{(nd_u)} = \frac{d_u}{Q'_{uu}(1)} > 0$;
- Si u est nul (non positif), alors $\lim_n p_{uu}^{(n)} = 0$.

Critère de positivité

Pour tout état u, on appelle période de u, et on note d_u, le pgcd des n tels que $p_{uu}^{(n)} > 0$. C'est aussi le pgcd des longueurs de cycles dans la composante fortement connexe contenant u. En particulier, si un sommet possède une boucle ($p_{uu} > 0$), tout sommet de sa composante fortement connexe a période 1.

Théorème :

- Si u est positif, alors $\lim_n p_{uu}^{(nd_u)} = \frac{d_u}{Q'_{uu}(1)} > 0$;
- Si u est nul (non positif), alors $\lim_n p_{uu}^{(n)} = 0$.

(admis)

Remarque: version séries génératrices…

C'est en fait un théorème sur les séries entières à coefficients positifs ou nuls :
Théorème :
Soit $Q(t) = \sum_{n\geq 1} q_n t^n$ une série à coefficients positifs ou nuls, avec $Q(1) = 1$ (rayon de convergence au moins 1), et telle que $q_n > 0$ pour n assez grand.
On pose $P(t) = 1/(1 - Q(t)) = \sum_{n\geq 0} p_n t^n$; les coefficients p_n satisfont

$$p_n = \delta_{0,n} + \sum_{1\leq k\leq n} q_k p_{n-k}$$

Alors, en posant $\tau = Q'(1) \leq +\infty$, on a

$$\lim_n p_n = \frac{1}{\tau}$$

Cas fini

Si l'espace d'états est fini, il n'y a pas de différence entre états récurrents et états positifs : tout état récurrent est automatiquement positif.

Preuve : $p_{uu}^{(n+m)} \geq p_{uv}^{(n)} p_{vu}^{(m)}$, donc si u est récurrent et nul, on a forcément $\lim_n p_{uv}^{(n)} = 0$ pour tout v dans la composante de u. En sommant sur v, on aurait $\lim_n \sum_v p_{uv}^{(n)} = 0$, ce qui est incompatible avec $\sum_v p_{uv}^{(n)} = 1$. Donc, tout état récurrent est positif.

Distribution limite?

Pour la chaîne partant de l'état u, la suite $(p_{uv}^{(n)})_{n\geq 0}$ a-t-elle une limite ?

Distribution limite?

Pour la chaîne partant de l'état u, la suite $(p_{uv}^{(n)})_{n\geq 0}$ a-t-elle une limite ?
Oui, mais en toute généralité, au sens de Cesaro seulement

$$\lim_n \frac{1}{n} \sum_{k\leq n} p_{uv}^{(k)} = \frac{Q_{uv}(1)}{Q'_{vv}(1)} = \frac{\mathbb{P}(T_{uv} < +\infty)}{\mathbb{E}(T_{vv})} = \pi_u(v)$$

Distribution limite?

Pour la chaîne partant de l'état u, la suite $(p_{uv}^{(n)})_{n\geq 0}$ a-t-elle une limite ?
Oui, mais en toute généralité, au sens de Cesaro seulement

$$\lim_n \frac{1}{n}\sum_{k\leq n} p_{uv}^{(k)} = \frac{Q_{uv}(1)}{Q'_{vv}(1)} = \frac{\mathbb{P}(T_{uv} < +\infty)}{\mathbb{E}(T_{vv})} = \pi_u(v)$$

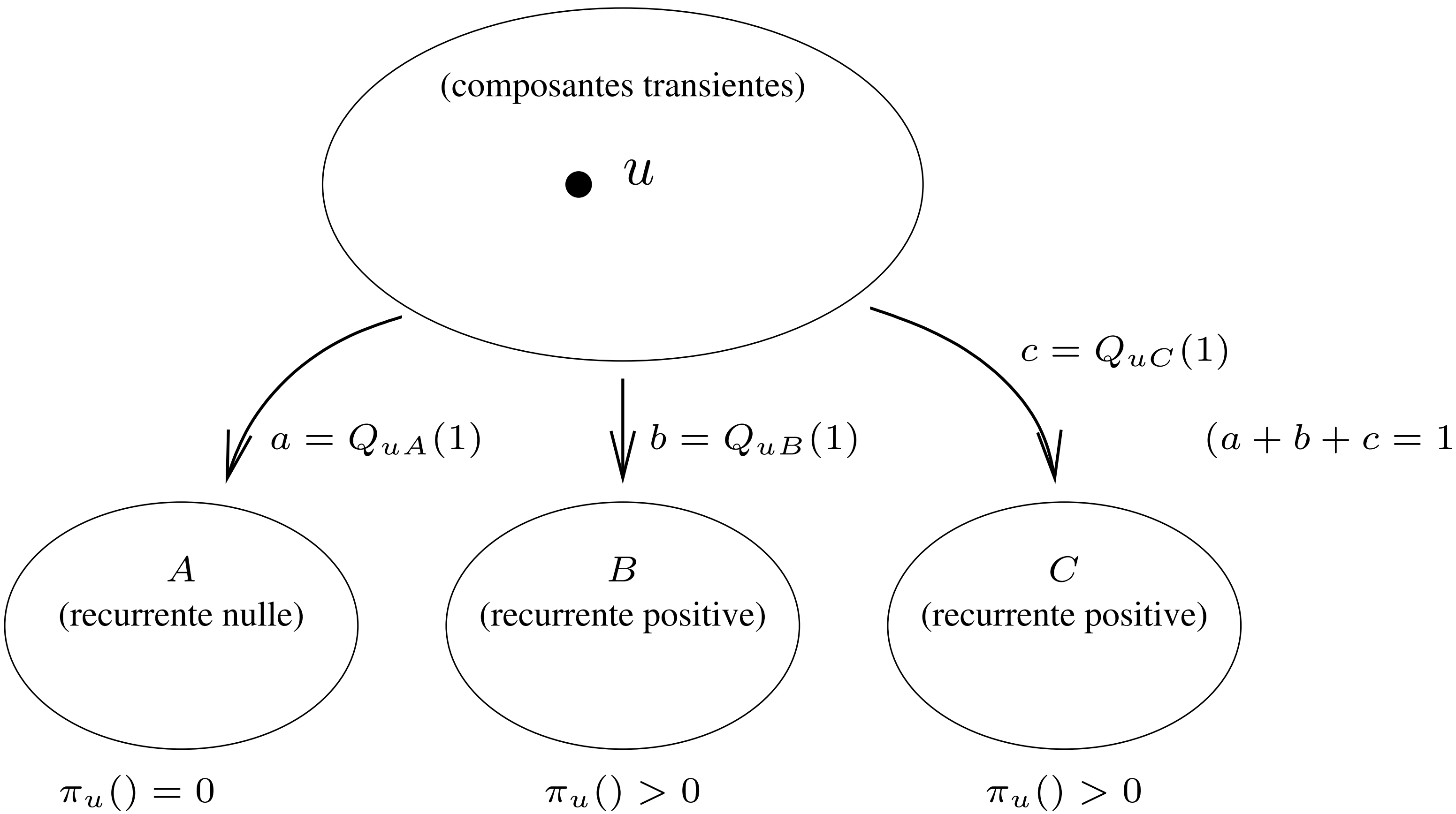
Idée de preuve :

$$\frac{1}{n}\sum_{k\leq n} p_{uv}^{(k)} = \frac{1}{n}\sum_{i\leq n} p_{vv}^{(i)} \sum_{j\leq n-i} q_{uv}^{(j)}$$

Convergence de Cesaro du théorème pour $p_{vv}^{(i)}$, convergence de $\sum_{j\leq n-i} q_{uv}^{(j)}$ vers $Q_{uv}(1)$, positivité. . .

Lois stationnaires?

Si la chaîne a des composantes récurrentes non positives, π_u peut ne pas définir une loi de probabilités sur l'espace d'états…



Et sinon?

π_u peut être identiquement nulle : si toutes les composantes récurrentes accessibles à partir de u sont nulles. Si toutes les composantes récurrentes sont nulles, il n'y a pas de loi de probabilités stationnaire.
Si π_u n'est pas identiquement nulle, $\pi_u / \sum_w \pi_u(w)$ est une loi de probabilités sur E, et elle est stationnaire.

$$\begin{aligned}
\frac{1}{n} p_{uv}^{(n)} &= \frac{n-1}{n} \sum_w \left(\frac{1}{n-1} p_{uw}^{(n-1)} \right) p_{wv} \\
\pi_u(v) &= \sum_w \pi_u(w) p_{wv}
\end{aligned}$$

La loi stationnaire n'est pas *a priori* unique : potentiellement, à chaque u correspond une π_u différente...

Cas d'unicité

Si la chaîne admet deux composantes positives distinctes $\mathcal{C}_1$ et $\mathcal{C}_2$, les lois stationnaires issues d'éléments des deux composantes sont forcément distinctes (l'une ne charge que $\mathcal{C}_1$, l'autre, que $\mathcal{C}_2$) : pour que la loi stationnaire soit unique, il est nécessaire qu'il n'y ait qu'une seule composante récurrente positive.

Cas d'unicité

Si la chaîne admet deux composantes positives distinctes $\mathcal{C}_1$ et $\mathcal{C}_2$, les lois stationnaires issues d'éléments des deux composantes sont forcément distinctes (l'une ne charge que $\mathcal{C}_1$, l'autre, que $\mathcal{C}_2$) : pour que la loi stationnaire soit unique, il est nécessaire qu'il n'y ait qu'une seule composante récurrente positive.
Réciproquement, c'est moins évident, mais vrai tout de même : s'il existe une unique composante récurrente positive, alors la loi stationnaire est unique et c'est celle obtenue comme limite de Cesaro en faisant partir la chaîne de n'importe lequel des états de cette composante.

Cas d'unicité

Si la chaîne admet deux composantes positives distinctes $\mathcal{C}_1$ et $\mathcal{C}_2$, les lois stationnaires issues d'éléments des deux composantes sont forcément distinctes (l'une ne charge que $\mathcal{C}_1$, l'autre, que $\mathcal{C}_2$) : pour que la loi stationnaire soit unique, il est nécessaire qu'il n'y ait qu'une seule composante récurrente positive.
Réciproquement, c'est moins évident, mais vrai tout de même : s'il existe une unique composante récurrente positive, alors la loi stationnaire est unique et c'est celle obtenue comme limite de Cesaro en faisant partir la chaîne de n'importe lequel des états de cette composante.
(Les lois stationnaires sont exactement les combinaisons convexes des lois obtenues sur chacune des composantes récurrentes positives)

Résumons. . .

- Il existe une loi stationnaire si et seulement si il existe une composante récurrente positive; si ce n'est pas le cas, toutes les probabilités de présence tendent vers 0.

Résumons. . .

- Il existe une loi stationnaire si et seulement si il existe une composante récurrente positive; si ce n'est pas le cas, toutes les probabilités de présence tendent vers 0.
- Les lois stationnaires ne chargent que les composantes récurrentes positives;

Résumons. . .

- Il existe une loi stationnaire si et seulement si il existe une composante récurrente positive; si ce n'est pas le cas, toutes les probabilités de présence tendent vers 0.
- Les lois stationnaires ne chargent que les composantes récurrentes positives;
- La loi stationnaire est unique, si et seulement si la composante récurrente positive l'est;

Résumons…

- Il existe une loi stationnaire si et seulement si il existe une composante récurrente positive; si ce n'est pas le cas, toutes les probabilités de présence tendent vers 0.
- Les lois stationnaires ne chargent que les composantes récurrentes positives;
- La loi stationnaire est unique, si et seulement si la composante récurrente positive l'est;
- Quand il y a une loi stationnaire, les probabilités de présence tendent au sens de Cesaro vers 0 ou vers une telle loi;

Résumons. . .

- Il existe une loi stationnaire si et seulement si il existe une composante récurrente positive; si ce n'est pas le cas, toutes les probabilités de présence tendent vers 0.
- Les lois stationnaires ne chargent que les composantes récurrentes positives;
- La loi stationnaire est unique, si et seulement si la composante récurrente positive l'est;
- Quand il y a une loi stationnaire, les probabilités de présence tendent au sens de Cesaro vers 0 ou vers une telle loi;
- Si les composantes positives chargées par la loi sont toutes de période 1, il y a convergence simple.

Théorème ergodique

Si la chaîne est irréductible, positive (l'unique composante est récurrente et positive) et apériodique, alors quelle que soit la distribution de X_0, toutes les probabilités de présence tendent vers l'unique loi stationnaire. Cette loi accorde un poids strictement positif à chaque état.

Théorème ergodique

Si la chaîne est irréductible, positive (l'unique composante est récurrente et positive) et apériodique, alors quelle que soit la distribution de X_0, toutes les probabilités de présence tendent vers l'unique loi stationnaire. Cette loi accorde un poids strictement positif à chaque état.
De plus, avec probabilité 1, pour tout état v, la proportion des instants entre 0 et N où la chaîne est dans l'état v tend vers $\pi(v)$ (pour presque toutes les trajectoires, la moyenne dans le temps est identique à la moyenne dans l'espace).

Génération aléatoire

On peut envisager de transformer une chaîne de Markov ergodique en générateur aléatoire d'états : si on dispose d'une chaîne ergodique, dont la distribution limite (stationnaire) est π (par exemple, uniforme), et que l'on sait simuler, on obtient un générateur aléatoire suivant une loi "presque égale" à π en partant d'un état X_0 quelconque, en simulant la chaîne pendant un temps N "assez grand", et en retournant X_N.

Génération aléatoire

On peut envisager de transformer une chaîne de Markov ergodique en générateur aléatoire d'états : si on dispose d'une chaîne ergodique, dont la distribution limite (stationnaire) est π (par exemple, uniforme), et que l'on sait simuler, on obtient un générateur aléatoire suivant une loi “presque égale” à π en partant d'un état X_0 quelconque, en simulant la chaîne pendant un temps N “assez grand”, et en retournant X_N.

- Comment détermine-t-on π ?

Génération aléatoire

On peut envisager de transformer une chaîne de Markov ergodique en générateur aléatoire d'états : si on dispose d'une chaîne ergodique, dont la distribution limite (stationnaire) est π (par exemple, uniforme), et que l'on sait simuler, on obtient un générateur aléatoire suivant une loi "presque égale" à π en partant d'un état X_0 quelconque, en simulant la chaîne pendant un temps N "assez grand", et en retournant X_N.

- Comment détermine-t-on π ?
- Ça veut dire quoi, "presque égale"?

Génération aléatoire

On peut envisager de transformer une chaîne de Markov ergodique en générateur aléatoire d'états : si on dispose d'une chaîne ergodique, dont la distribution limite (stationnaire) est π (par exemple, uniforme), et que l'on sait simuler, on obtient un générateur aléatoire suivant une loi "presque égale" à π en partant d'un état X_0 quelconque, en simulant la chaîne pendant un temps N "assez grand", et en retournant X_N.

- Comment détermine-t-on π ?
- Ça veut dire quoi, "presque égale"?
- C'est quoi, "assez grand"?

Génération aléatoire

On peut envisager de transformer une chaîne de Markov ergodique en générateur aléatoire d'états : si on dispose d'une chaîne ergodique, dont la distribution limite (stationnaire) est π (par exemple, uniforme), et que l'on sait simuler, on obtient un générateur aléatoire suivant une loi "presque égale" à π en partant d'un état X_0 quelconque, en simulant la chaîne pendant un temps N "assez grand", et en retournant X_N.

- Comment détermine-t-on π ?
- Ça veut dire quoi, "presque égale"? (distance)
- C'est quoi, "assez grand"?

Génération aléatoire

On peut envisager de transformer une chaîne de Markov ergodique en générateur aléatoire d'états : si on dispose d'une chaîne ergodique, dont la distribution limite (stationnaire) est π (par exemple, uniforme), et que l'on sait simuler, on obtient un générateur aléatoire suivant une loi "presque égale" à π en partant d'un état X_0 quelconque, en simulant la chaîne pendant un temps N "assez grand", et en retournant X_N.

- Comment détermine-t-on π ?
- Ça veut dire quoi, "presque égale"? (distance)
- C'est quoi, "assez grand"? (temps de mélange)