Méthodes de Monte–Carlo

EISC-211

Xavier Gendre & Jérôme Morio

ISAE & ONERA

2018–2019

Très rapide aperçu des chaînes de Markov

__But__ : Algorithme de Metropolis–Hastings

Définition

Chaîne de Markov $(X_t, t \geq 1)$

- **Système dynamique aléatoire**
- **Innovation indépendante du passé**

$$X_{t+1} = F(X_t, U_t), \ (U_t) \ \textbf{i.i.d.}$$

Utilité

Modélisation de nombreux systèmes physiques

- L'évolution future ne dépend que de l'état actuel du système et pas de comment il y est arrivé
- Marchés financiers, séquences d'ADN, évolution de population, réseaux de communication, formation de cristaux, ...

Algorithmes probabilistes

- Algorithmique
- Optimisation : recuit simulé
- Page Rank
- Simulation

Propriété de Markov

Définition

Un processus $(X_t, t \geq 1)$ satisfait la propriété de Markov si le futur est indépendant du passé conditionnellement au présent, i.e., si

$$\mathbb{P}\left(X_{t+1} = x_{t+1} \mid X_t = x_t, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0\right) \\ = \mathbb{P}\left(X_{t+1} = x_{t+1} \mid X_t = x_t\right)$$

Analogie déterministe :

$\dot{x}(t) = F(x(t))$ **et non** $\dot{x}(t) = F(x(s), s \leq t)$

Théorème

$(X_t, t \geq 1)$ **satisfait la propriété de Markov si et seulement on peut écrire**

$$X_{t+1} = F(X_t, U_t),\ \ (U_t)\ \ \textbf{i.i.d.}$$

Propriété de Markov

Matrice de transition

La dynamique d'un processus de Markov (X_t) est donc caractérisée par la matrice de transition :

$$P = (p_{xy})_{x,y\in\mathfrak{X}} \quad \textbf{avec} \quad p_{xy} = \mathbb{P}(X_1 = y \mid X_0 = x)$$

Convergence à l'équilibre

Rappel

(X_t) **est une suite de variables aléatoires : elle peut donc converger en loi.**

Théorème

Si

$$\boxed{\forall x, y \in \mathfrak{X} : \pi_x p_{xy} = \pi_y p_{yx}}$$

alors $X_t \xrightarrow[t \to +\infty]{L} \pi$**, i.e.,** $\mathbb{P}(X_t = x) \xrightarrow[t \to +\infty]{} \pi_x$ **pour tout** $x \in \mathfrak{X}$**.**

Le cas i.i.d.

Cas particulier

Si $p_{xy} = p_y$, i.e., l'état suivant ne dépend pas de la position courante, alors (X_t) est i.i.d. distribuée selon $p = (p_x)_{x \in \mathfrak{X}}$.

- **Loi des grands nombres**
- **Théorème central limite**
- **IIDMC**

Ces deux résultats restent vrais dans le cas général !

- **IIDMC $\rightarrow$ MCMC**

Théorèmes de convergence

Théorème ergodique

Si $X_t \xrightarrow[t\to+\infty]{L} f$ **+ (...), alors**

$$\Phi_T = \frac{1}{T}\sum_{t=1}^{T}\phi(X_t) \xrightarrow[T\to+\infty]{p.s.} \mathbb{E}_f(\phi(X_1)) = \int \phi f.$$

Théorème central limite

Si $X_t \xrightarrow[t\to+\infty]{L} f$ **+ (...), alors**

$$\sqrt{T}\left(\Phi_T - \int \phi f\right) \xrightarrow[T\to+\infty]{L} N_f$$

où N_f **suit une loi normale centrée de variance**

$$\mathbb{V}\mathsf{ar}_f(\phi(X_1)) + 2\sum_{t\geq 2}\mathbb{C}\mathsf{ov}_\pi(\phi(X_1), \phi(X_t))$$

Remarque

Cas i.i.d. $\Rightarrow$ $\mathbb{C}\mathsf{ov}_\pi(\phi(X_1), \phi(X_t)) = 0$

Théorèmes de convergence

Remarque

- (X_t) **i.i.d.** $\Rightarrow$ $(\phi(X_t))$ **i.i.d.**
- (X_t) **MC** $\not\Rightarrow$ $(\phi(X_t))$ **MC !**

Résumé

En résumé :

Si $\pi_x p_{xy} = \pi_y p_{yx} (\star)$, alors on a la loi des grands nombres et le TCL.

Algorithme de Metropolis–Hastings

- **Etant donné une distribution (densité) cible π, construction systématique d'une chaîne de Markov (P) qui satisfait $(\star)$**
- **Applications inombrables : biologie, physique, chimie, finances, fiabilité, traitement du signal, optimisation, EDP, vision par ordinateur, . . .**
- **Très souvent la seule méthode disponible**
- **(A des défauts)**