Université Libre de Bruxelles

INFO-F311
Intelligence Artificielle

Synthèse IA

Étudiants :
Rayan Contuliano Bravo

Enseignants :
T. Lenaerts

25 décembre 2023

Contents

1 Modèle de Markov

Définition 1.1. Modèles de Markov

Les **Modèles de Markov** sont une sorte de réseaux bayésiens qui décrivent l'évolution d'un système dans le temps. Ils sont composés de variables aléatoires $X_1, X_2, \ldots, X_n$ qui représentent l'état du système à différents moments dans le temps. Les variables aléatoires sont liées par des probabilités de transition $P(X_t|X_{t-1})$ qui décrivent la probabilité de passer d'un état à l'autre, comment l'état du système évolue dans le temps.

Ces modèles sont utiles pour prédire l'état d'un système à un moment donné.

Théorème 1.1. Processus Markovien

Soit $X_1, X_2, \ldots, X_n$ une séquence de variables aléatoires. Si la propriété de Markov est respectée, alors

$$P(X_t|X_{t-1}, X_{t-2}, \ldots, X_1) = P(X_t|X_{t-1})$$

Cela signifie que **le future** est **indépendant** du **passé**, étant donné le **présent**.

Théorème 1.2. Proccessus Stationnaire

Soit $X_1, X_2, \ldots, X_n$ une séquence de variables aléatoires. Si la propriété de Markov est respectée, alors

$$P(X_t|X_{t-1}) = P(X_{t+1}|X_t)$$

Cela signifie que la probabilité de transition est la même pour tous les instants. On dit que le modèle est **stationnaire**.

Remarque: Lorsque l'instant t est indépendant des instants $t-2, t-3, \ldots, 1$, sachant $t-1$, on dit que le modèle est d'ordre 1. Il este possible d'avoir des modèles d'ordre supérieur, dans lequel l'instant t dépends des k instants précédents.

Pour calculer la probabilité de distribution conjointe d'un modèle de Markov, il faut utiliser la propriété de Markov et la règle de la chaîne.

$$\begin{aligned} P(X_1, X_2, \ldots, X_n) &= P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)\ldots P(X_n|X_{n-1}, \ldots, X_1) \\ &= P(X_1)\prod_{t=2}^{n} P(X_t|X_{t-1}) \end{aligned}$$

1.1 Le modele qu'on ne trouve pas

La différence avec un modèle de Markov, c'est que un modèle de Markov caché est un modèle dans lequel après chaque état, une observation est générée. On peut alors utiliser les observations pour inférer l'état du système.

Nous supposons également que l'observation est générée **uniquement par** l'état courant du système.

$$P(E_t|X_{1:t}, E_{1:t-1}) = P(E_t|X_t)$$

Sachant que l'état du système est connu, l'observation est indépendante des observations précédentes ainsi que des états précédents.

Exemple: Un gardien de sécurité passe un mois dans un édifice sous-terrain, sans sortir. Chaque jour, son directeur arrive avec ou sans parapluie. Le gardien veut inférer la possibilite qu'il ait plu ou non en fonction de la sequence des observations du parapluie.

Voici le modèle de Markov pour ce problème.

— **Variables** : $X_t = R_t$ qui représentre *Rain* et $E_t = U_t$ qui représente *Umbrella*.

— Dépendances entre les variables, càd Modèle de Markov :

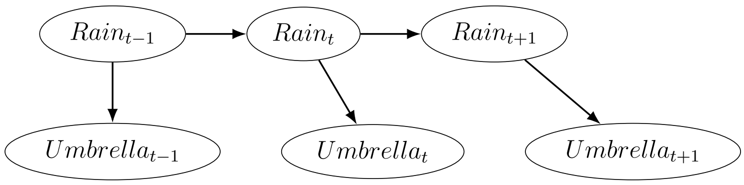

Figure 1.1 – Modèle de Markov

— Modèle de transitions : $P(R_t|R_{t-1})$ et $P(U_t|R_t)$ *Stationnaire*

Nous voyons bien que l'état du système à un moment donné ne dépend que de l'état précédent. Et que les observations sont générées uniquement par l'état courant du système.

Pour calculuer la probabilité de distribution conjointe d'un modèle de Markov caché, il faut utiliser la propriété de Markov et la règle de la chaîne.

$$\begin{aligned} P(X_1, X_2, \ldots, X_n, E_n) &= P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)\ldots P(X_n|X_{n-1}, \ldots, X_1)P(E_n|X_n) \\ &= P(X_1)\prod_{t=2}^{n} P(X_t|X_{t-1})P(E_t|X_t) \end{aligned}$$

1.2 Inférence

Il y a plusieurs problèmes d'inférence que l'on peut résoudre avec les modèles de Markov cachés.

— **Filtering** : Calculer la distribution de probabilité de l'état courant du système, sachant toutes les observations jusqu'à maintenant.
$$P(X_t|E_{1:t})$$

— **Prediction** : Calculer la distribution de probabilité de l'état futur du système, sachant toutes les observations jusqu'à maintenant. Sans tenir compte des observations futures.
$$P(X_{t+k}|E_{1:t})$$

— **Smoothing** : Calculer la distribution de probabilité de l'état passé du système, sachant toutes les observations jusqu'à maintenant. Meilleure estimation de l'état passé. *Learning*
$$P(X_{t-k}|E_{1:t})$$

— **Most Likely Explanation** : Calculer la séquence d'états **la plus probable**, sachant toutes les observations jusqu'à maintenant.
$$\arg\max_{X_{1:t}} P(X_{1:t}|E_{1:t}) = \arg\max_{X_{1:t}} P(X_{1:t}, E_{1:t})/P(E_{1:t}) = \arg\max_{X_{1:t}} P(X_{1:t}, E_{1:t})$$

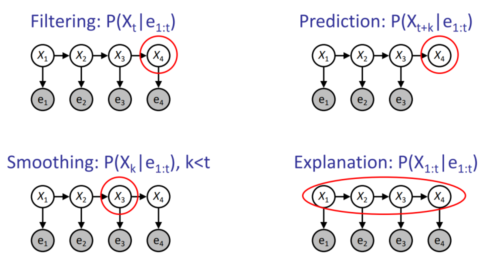

Figure 1.2 – Inférence dans un modèle de Markov

1.3 Chaines de Markov

Les chaines de Markov sont un cas particulier des modèles de Markov. Elles sont composées d'un ensemble d'états $S = s_1, s_2, \ldots, s_n$ et d'une matrice de transition T qui décrit la probabilité de passer d'un état à l'autre.

$$T_{ij} = P(X_t = s_j | X_{t-1} = s_i)$$

(même définition que pour les modèles de Markov, c'est juste moins général)