Algorytmy robotyki mobilnej
Podstawy matematyczne – przypomnienie

Janusz Jakubiak

Katedra Cybernetyki i Robotyki

2020/2021

Politechnika Wrocławska

HR EXCELLENCE IN RESEARCH

Informacja o prawach autorskich

Elementy teorii prawdopodobieństwa

- Aksjomaty
  - $0 \leq p(X) \leq 1$
  - $p(Prawda) = 1 \qquad p(Falsz) = 0$
  - $p(X \cup Y) = p(X) + p(Y) - p(X \cap Y)$
- Dyskretne zmienne losowe
  - X jest zmienną losową przyjmującą wartości ze zbioru$\{x_1, x_2, \ldots, x_n\}$
  - $p(X = x_i)$ oznacza prawdopodobieństwo, że zmienna X przyjmuje wartość x_i
- Ciągłe zmienne losowe
  - X jest zmienną losową przyjmującą wartości z $\mathcal{R}$
  - $p(X = x)$ (lub $p(x)$) oznacza gęstość prawdopodobieństwa w x

Prawdopodobieństwo warunkowe

- Prawdopodobieństwo łączne $P(x, y) = P(X = x \wedge Y = y)$
- W przypadku zdarzeń niezależnych $P(x, y) = P(x)P(y)$
- Prawdopodobieństwo warunkowe $P(x|y)$ jest prawdopodobieństwem, że $X = x$ jeśli $Y = y$
- $P(x, y) = P(x|y)P(y)$
- Jeśli X i Y są niezależne, to $P(x|y) = P(x)$

Własności

	dyskretne	ciągłe
prawd. łączne	$\sum_x P(x) = 1$	$\int p(x)dx = 1$
prawd. całkowite	$P(x) = \sum_y P(x\|y)P(y)$	$p(x) = \int p(x\|y)p(y)dy$
brzegowa f. prawd.	$P(x) = \sum_y P(x,y)$	$p(x) = \int p(x,y)dy$

5

Notacja

- x – stan rzeczywisty układu (R^n), $\bar{x}$ – estymata stanu,
- z – obserwacja (R^m),
- $\mathcal{N}(\mu, \Sigma)$ – wielowymiarowy rozkład normalny ze średnią μ i macierzą kowariancji Σ (parametrami: momenty)

$$\mathcal{N}(\mu, \Sigma) \sim p(X) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$

- Wzór Bayesa

$$p(X|Z) = \frac{p(Z|X)p(X)}{p(Z)} = \eta p(Z|X)p(X),$$

gdzie η jest współczynnikiem normalizującym.

6

Notacja (2)

Zmiany momentów dystrybucji gaussowskich przy przekształceniach

- liniowym ($Y = AX + B$)

$$p(X) = \mathcal{N}(\mu, \Sigma) \quad \rightarrow \quad p(Y) = \mathcal{N}(A\mu + B, A\Sigma A^T)$$

- sumie ($Y = X_1 + X_2$)

$$p(X_1) = \mathcal{N}(\mu_1, \Sigma_1), p(X_2) = \mathcal{N}(\mu_2, \Sigma_2), \quad \rightarrow \quad p(Y) = \mathcal{N}(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)$$

- iloczyn

$$p(X_1)p(X_2) = \mathcal{N}(\mu_1, \Sigma_1)\mathcal{N}(\mu_2, \Sigma_2) \quad \rightarrow \quad \mathcal{N}(\mu_Y, \Sigma_Y),$$

$$\Sigma_Y = \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1}$$
$$\mu_Y = \Sigma_Y \Sigma_1^{-1}\mu_1 + \Sigma_Y \Sigma_2^{-1}\mu_2$$

7

Przykłady – wzór Bayesa

8

Proces Markowa (jednorodny, stacjonarny)

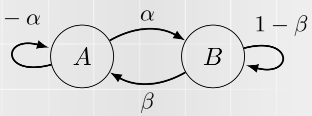
Proces strochastyczny nad skończoną przestrzeni stanów, w którym bieżący stan zależy wyłącznie od stanu poprzedniego oraz stałego w czasie prawdopodobieństwa przejścia.

Formalnie:

$$P(X_{k+1}|X_k X_{k-1} \dots X_1 X_0) = P(X_{k+1}|X_k)$$

lub:
proces bez pamięci

Przykładowy proces dwustanowy:



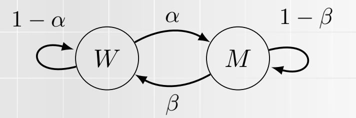
Przykład 1 – łańcuch Markowa

Każdego dnia robot może działać, lub mieć awarię. Po dniu poprawnej pracy prawdopodobieństwo poprawnej pracy kolejnego dnia wynosi 0.7, a jeśli danego dnia robot ma awarię prawdopodobieństwo, że następnego zostanie naprawiony i będzie pracował prawidłowo wynosi 0.9.
Jeśli wiemy, że w dniu 0 pracuje, jakie jest prawdopodobieństwo, że

- pracuje przez 3 kolejne dni?
- pracuje przez dokładnie 2 z 3 dni?

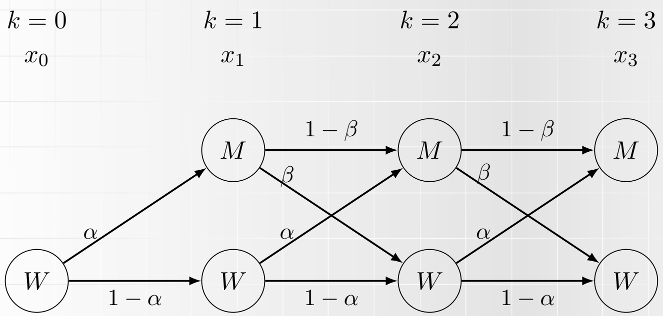
$$p(W|W) = 1 - \alpha = 0.7$$
$$p(W|M) = \beta = 0.9$$

$S = \{W, M\}$



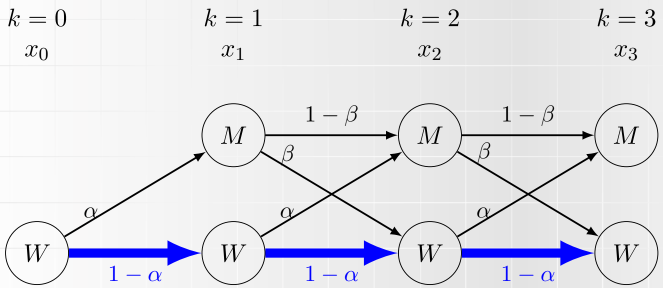
$x_k \backslash x_{k+1}$	W	M
W	0.7	0.3
M	0.9	0.1

Przykład 1 – łańcuch Markowa



Przykład 1 – łańcuch Markowa

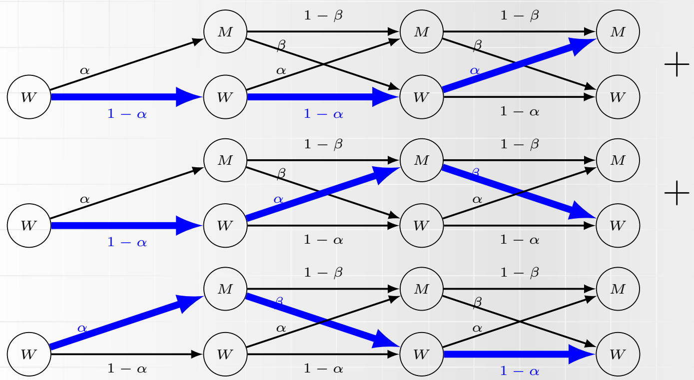
…prawdopodobieństwo pracy przez 3 kolejne dni?



$$p(X = \{WWW\}) =$$
$$p(x_3 = W|x_2 = W)p(x_2 = W|x_2 = W)p(x_1 = W|x_0 = W) =$$
$$0.7 \cdot 0.7 \cdot 0.7 = 0.343$$

Przykład 1 – łańcuch Markowa

…prawdoodobieństwo pracy przez 2 z 3 dni?



Modele Markowa

- Łańcuch Markowa
- Ukryty model Markowa (HMM)

  Proces Markowa, w którym stan nie jest obserwowalny bezpośrednio, ale poprzez funkcję wyjścia.
- Proces decyzyjny Markowa (MDP)

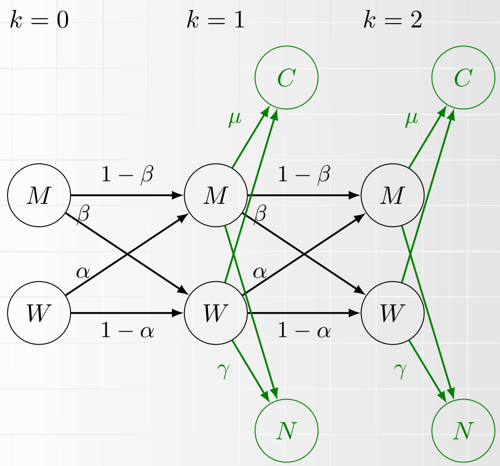
  Łańcuch Markowa z dodanym sterowaniem (akcją) – nowy stan zależy od pary (stan poprzedni, akcja).
- Częściowo obserwowalny proces decyzyjny Markowa (POMDP)

  MDP z funkcją wyjścia.

HMM – opis

- Przestrzeń stanów $S = \{S_1, \ldots, S_n\}$
- Przestrzeń obserwacji $O = \{O_1, \ldots, O_m\}$
- Ciąg stanów $X = \{x_0, x_1, \ldots\}, \quad x_i \in S$
- Ciąg obserwacji $Z = \{z_1, z_2 \ldots\}, \quad z_i \in O$
- Prawdopodobieństwo początkowe $p_{0i} = P(X_0 = S_i), \quad i = 1, \ldots n$ (lub stan początkowy $x_0 \in S$)
- Macierz przejść $A = \{A_{ij} = P(X_{k+1} = S_j | X_k = S_i)\}$
  uwaga: $\sum_j A_{ij} = 1$
- Macierz obserwacji $C = \{C_{ij} = P(Z_k = O_j | X_k = S_i)\}$
  uwaga: $\sum_j C_{ij} = 1$

Przykład 2 – HMM



$x_k \backslash x_{k+1}$	W	M
W	$1-\alpha$	α
M	β	$1-\beta$

$x_k \backslash z_k$	C	N
W	$1-\gamma$	γ
M	μ	$1-\mu$

HMM – 3 główne zadania

- Jakie jest prawdopodobieństwo danego ciągu obserwacji Z?

  algorytm „forward-backward"
- Jaki jest najbardziej prawdopodobny ciąg stanów, który daje dany ciąg obserwacji?

  Algorytm Viterbiego
- Jakie są wartości parametrów macierzy przejścia i obserwacji?

  Algorytm Bauma-Welcha

Algorytm Viterbiego

Szukane

$$x^\star = \arg\max_X P(X|Z)$$

Złożoność obliczeniowa: N^K

Idea rozwiązania: jeśli wiemy, że $\bar{X}^\star = \{x_0 x_1 \dots x_k\}$ jest najbardziej prawdopodobnym ciągiem stanów dla obserwacji $\bar{Z} = \{z_1 \dots z_k\}$, to po dodaniu kolejnej obserwacji $Z = \{z_1 \dots z_k z_{k+1}\}$, najbardziej prawdopodobny ciąg stanów $X^\star = \{x_0 x_1 \dots x_k x_{k+1}\}$ zawiera $\bar{X}^\star$ ($X^\star = \{\bar{X}^\star, x_{k+1}\}$).

(por. zasada optymalności Bellmana, programowanie dynamiczne)

Algorytm Viterbiego

Algorytm korzysta z prawdopodobieństwa warunkowego

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z)}$$

w procedurze rekurencyjnej: jeśli

$$\delta_{kj} = \max_{\bar{X}} P(\bar{X}|Z)$$

jest najbardziej prawdopodobną sekwencją, to najlepsze rozwiązanie dla kolejnego kroku może być wyliczone jako

$$\delta_{k+1,j} = \max_i \{\delta_{ki} A_{ij}\} B_{j,k+1}$$

Pytania podsumowujące

1. Czym jest i jak korzystać z tw. Bayesa?
2. Jakie własności mają operacje na zmiennych o rozkładzie normalnym?
3. Jakie własności i zastosowania mają modele Markowa?