**Uniwersytet Wrocławski**
**Wydział Matematyki i Informatyki**
**Instytut Matematyczny**
*specjalność: Analiza Danych*

***Klaudia Balcer***

# FlowHMM

Praca magisterska
napisana pod kierunkiem
dr. hab. Pawła Lorka

Wrocław 2023

## Pytania

1. Praca może być po angielsku?
2. Czy taka strukutra pracy będzie OK?
3. Czy rozpisywać dokładnie przykłady (estymację parametrów dla różnych emisji)?

# Contents

# Abstract

# 1 Introduction

- Minimalny wstęp (ta praca coś tam coś tam)
- Czym (z grubsza) jest HMM?
- Struktura dokumentu

# 2 Standard Solution

## 2.1 Markov Chains

- Markov Chain
- time-homogenous Markov Chains
- irreproducuble Markov Chain
- aperiodic markov chain
- ergodic Markov chain

$$\mathcal{S} = \{s_1, \ldots, s_N\}$$

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix}$$

$$A(i,j) = \mathbb{P}(X_{k+1} = s_j | X_k = s_i)$$

$$\pi = (\pi_1, \pi_2, \ldots, \pi_N)$$

$$\{X_k\}_{k \geq 0}$$

- ergodic Markov Chain

Markov Assumption:

$$\mathbb{P}(X_{k+1} = x_{k+1} | X_{0:k} = x_{0:k}) = \mathbb{P}(X_{k+1} = x_{k+1} | X_k = x_k)$$

## 2.2 Hidden Markov Model

$$\mathcal{V}$$

- observation space (discrete or continuous)

$$\{Y_k\}_{k\geq 0}$$

sequence of random variables taking values from $\mathcal{V}$

Independence Assumption:

$$\mathbb{P}(Y_{k+1} = y_{k+1}|Y_{0:k} = y_{0:k}, X_{0:k+1} = x_{0:k+1}) = \mathbb{P}(Y_{k+1} = y_{k+1}|X_{k+1} = x_{k+1}) =: p(y_{k+1}|x_{k+1})$$

#### 2.2.1 Example 1: cathegorical HMM

#### 2.2.2 Example 2: Gaussian HMM

#### 2.2.3 Example 3: Gausian Mixtuure Model HMM

## 2.3 Parameter estimation

#### 2.3.1 Forward Algorithm

Forward probability:

$$\alpha_k(i) = \mathbb{P}(Y_{1:k} = y_{1:k}, X_k = s_i)$$

Recursive formula:

$$\alpha_1(i) = \pi_i \mathbb{P}(Y_1 = y_1|X_1 = s_i)$$
$$\alpha_k(i) = \sum_{j=1}^{n} \alpha_{k-1}(j) A(j,i) \mathbb{P}(Y_k = y_k|X_k = s_i)$$
$$\mathbb{P}(Y_{1:T} = y_{1:T}) = \sum_{i=1}^{n} \alpha_T(i)$$

#### 2.3.2 Backward Algorithm

Backward probability:

$$\beta_k(i) = \mathbb{P}(Y_{k+1:T} = y_{k+1:T}|X_k = s_i)$$

Recursive formula:

$$\beta_T(i) = 1$$
$$\beta_k(i) = \sum_{j=1}^{n} A(i,j) \mathbb{P}(Y_k = y_k|X_k = s_j) \beta_{k+1}(j)$$
$$\mathbb{P}(Y_{1:T} = y_{1:T}) = \sum_{j=1}^{N} \pi_j \mathbb{P}(Y_1 = y_1|X_1 = s_j) \beta_1(j)$$

#### 2.3.3 Forward-Backward Algorithm

$$\xi_k(i,j) := \mathbb{P}(X_{k+1} = s_j, X_k = s_i) \tag{3}$$

$$\gamma_k(i) := \mathbb{P}(X_k = s_i) \tag{4}$$

Expectations:

$$\gamma_k(i) = \mathbb{P}(X_k = s_i) = \frac{\mathbb{P}(X_k = s_i, Y_{0:T} = y_{0:T})}{\mathbb{P}(Y_{0:T} = y_{0:T})} = \frac{\alpha_k(i)\beta_k(i)}{\sum_{j=1}^{n} \alpha_k(j)\beta_k(j)}$$

$$\xi_k(i,j) = \mathbb{P}(X_{k+1} = s_j, X_k = s_i) = \frac{\alpha_k(i)A(i,j)\mathbb{P}(Y_{k+1} = y_{k+1}|X_{k+1} = s_j)\beta_{k+1}(j)}{\sum_{j=1}^{n} \alpha_k(j)\beta_k(j)}$$

Maximization: depends on distribution
**Examples?? (like previously)**

#### 2.3.4 Viterbi Algorithm

### 2.4 Limitations

#### 2.4.1 Related work - HMM extensions

- Różne rozszerzenia HMMów

## 3 Co-occurence based learning schema

### 3.1 Discretization

#### 3.1.1 Select points at random

#### 3.1.2 Quasi random

#### 3.1.3 unifromly by coordinates

#### 3.1.4 latin hypercude

## 4 Related work - HMM extensions

- Różne rozszerzenia HMMów

# 5 Moder approach: Flow HMM

## 5.1 Flow NN

## 5.2 Combined FlowHMM

## 5.3 Learning

### 5.3.1 Baum-Welch - adapted

### 5.3.2 Cooc-based learning - adapted

# 6 Experiments

## 6.1 Becnhmarks

## 6.2 Results

# 7 Conclusions