Sequence Modelling and Monte Carlo Methods

Hanchen Wang

Ph.D Candidate in
Engineering, University of Cambridge

March 11, 2020

Q1 Markov Models: fitting bi-gram models

(a) note that it is not HMM, thus we use k = 0, 1, 2 to represent the three states (A, B, C)

- the joint distribution of observation $y_{1:T}$ given initial state π^0 and transition matrix T is :

$$p\left(y_{1:T}|\pi^0, T\right) = \underbrace{\left(\prod_k \pi_k^0 \mathbb{1}(y_1 = k)\right)}_{\text{initial state}} \left(\prod_{t=2,k=1,l=1}^{T,K,K} T_{kl}^{\mathbb{1}(y_t=k, y_{t-1}=l)}\right) \underbrace{\begin{bmatrix} f & \gamma & w/2 \\ 0 & f & h/2 \\ 0 & 0 & 1 \end{bmatrix}}_{\substack{\text{intrinsics} \\ [\quad \mathbf{K} \quad]}}$$

- Bernouli Distribution Squeeze on the boundry(Q3)

Regression

- ML methods for regression find statistical dependencies not causal relationships:
  - for intro to 'causal', check Larry Wasserman's note:
    `http://www.stat.cmu.edu/~larry/=sml/Causation.pdf`
- Q6: noise can set as 'heavy tail distribution'[1]

[1] `https://en.wikipedia.org/wiki/Heavy-tailed_distribution`

Bayesian Linear Regression

- ML methods for regression find statistical dependencies not causal relationships:
  - for intro to 'causal', check Larry Wasserman's note:
    `http://www.stat.cmu.edu/~larry/=sml/Causation.pdf`
- Q6: noise can set as 'heavy tail distribution'[2]

[2] `https://en.wikipedia.org/wiki/Heavy-tailed_distribution`