Hidden Markov Models

Leonid Zhukov

School of Applied Mathematics and Information Science
National Research University Higher School of Economics

07.11.2013

Markov chain parameter estimation

For Markov chain transition matrix estimation

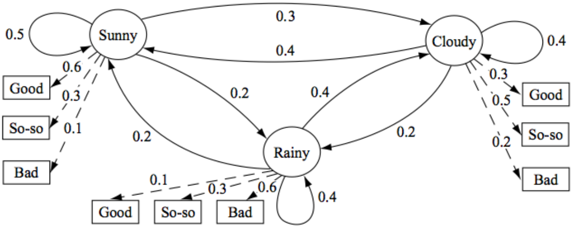
$$P(X_1, ..X_n|\theta) = P(X_1)\prod_i \prod_j p_{ij}^{n_{ij}} = L(p)$$

Maximum likelihood MLE estimate:

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$$

n_{ij} - number of times in sequence X_J followed X_i, i.e. transition $i \to j$

Hidden Markov Model



Hidden Markov Model

HMM formal definition

- Desrete states
- Observable signals
- Transion probabilities matrix $A^{N\times N}$, $A_{ij} = P(q_{t+1} = S_j | q_t = S_i)$
- Emmision probabilities matrix $B^{M\times N}$, $B_{ij} = b_i(O_j) = P(O_j | q_t = S_i)$
- Initial states vector π, $\pi_i = P(X_1 = i)$
- HMM Model $\lambda = (A, B, \pi)$

Three fundamental problems in HMM

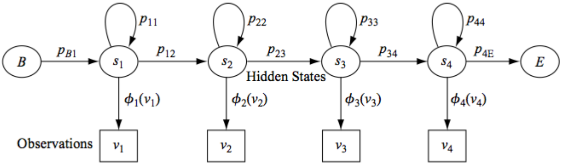
1. The Evaluation problem.
   Given:
   - Observable sequence $O = O_1 O_2 O_3 ... O_T$
   - model $\lambda = (A, B, \pi)$

   Find: $P(O|\lambda)$
2. The Decoding problem.
   Given:
   - Observable sequence $O = O_1 O_2 O_3 ... O_T$
   - model $\lambda = (A, B, \pi)$

   Find: $Q^* = q_1 q_2 q_3 .. q_T = \arg\max_Q P(Q|O, \lambda)$
3. The Learning problem (training).
   Given:
   - Observable sequence $O = O_1 O_2 O_3 ... O_T$

   Find: $\lambda^* = \arg\max_\lambda P(O|\lambda)$

Left-to-right HMM



References

- A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Lawrence R. Rabiner. Proc of IEEE, Vol 77, N 2, 1989, pp 257-286