CASE STUDY: Maximization of Log-Lokelihood in Hidden Markov Model (hmm_discrete, hmm_normal, linear, linearmulti)

Background

Approximation of set of random data by Hidden Markov Model may be a fruitful approach for many applications.

This case study considers two variants of Hidden Markov Model. One with discrete distributions of observations and other with normal distributions of observations. Correspondently two Problem statements for maximization of Log-Lokelihood function in Hidden Markov Model are shown.

For maximization type of problem PSG uses an expectation modification (EM) procedure in form of Baum–Welch algorithm to find good initial point.

hmm_discrete and *hmm_normal* functions report probabilities of initial states, transition probabilities and probabilities of observations or parameters of normal distributions. Additionally they report Viterbi states vector.

References

- Lawrence R. Rabiner (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of IEEE, 77-2, p. 267-295. http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/tutorial%20on%20hmm%20and%20applications.pdf

Notations

$\bar{O}$ = set of possible observations ($\bar{O} = \{O_1, \ldots, O_m\}$ for discrete distribution);

m = number of possible observations for discrete distribution;

T = number of consequence time moments;

$o_1, \ldots, o_T$ = sequence of observations, $o_t \in \bar{O}$;

n = number of hidden states to be considered in a model;

$\vec{x}$ = vector of decision variables consisting of $\vec{p}$, A and B for discrete distributed observations or

$\vec{x}$ = vector of decision variables consisting of $\vec{p}$, A and P for normal distributed observations,

where

$\vec{p} = (p_1, \ldots, p_n)$ = vector of probabilities of initial states such that $\sum_{i=1}^{n} p_i = 1$,

A = transition matrix with probabilities such that $\sum_{j=1}^{n} a_{ij} = 1$ for every $i = 1, \ldots, n$,

B = observations probabilities matrix with probabilities such that $\sum_{l=1}^{m} b_{il} = 1$ for every $i = 1, \ldots, n$,

P = matrix of parameters μ_i, σ_i of normal distributions $N(\mu_i, \sigma_i^2)$ for every $i = 1, \ldots, n$.

For discrete distributions

D_t = diagonal matrix with probabilities $d_{ii}^t = b_{il}, i = 1, \dots, n,\ t = 1, \dots, T$, where $o_t = O_l$.

For normal distributions

D_t = diagonal matrix with densities of probabilities $d_{ii}^t = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{(o_t-\mu_i)^2}{2\sigma_i^2}}$, $i = 1, \dots, n,\ t = 1, \dots, T$.

For both cases

$\boldsymbol{hmm\_function}(\vec{x}, \vec{o}) = ln\left(\vec{p}^T D_1 \left(\prod_{t=2}^{T} A\, D_t\right)\vec{1}\right)$ = Log-likelihood function for Hidden Markov Model.

$\vec{1} = (1, \dots, 1)$ is a vector with n ones.

Optimization Problem 1

maximizing Log-likelihood function for Hidden Markov Model with discrete distributions of observations

$$\max_{\vec{x}} \boldsymbol{hmm\_discrete}(\vec{x}, \vec{o})$$

subject to

constraints on probabilities of initial states

$$\boldsymbol{linear}(\vec{p}) = 1,$$

constraints on probabilities of transitions in every states

$$\boldsymbol{linearmulti}(A) = 1,$$

constraints on probabilities of observations in every state

$$\boldsymbol{linearmulti}(B) = 1,$$

box on variables

$$p_i \geq 0,\ \ a_{ij} \geq 0,\ \ b_{il} \geq 0,\ \ i = 1, \dots, n, j = 1, \dots, n, l = 1, \dots, m.$$

***Remark*:** variables for *hmm_discrete* function and constraints on probabilities are generated by PSG automatically, so user should not define these variables and constraints in Problem statement.

Optimization Problem 2

maximizing Log-likelihood function for Hidden Markov Model with normal distributions of observations

$$\max_{\vec{x}} \boldsymbol{hmm\_normal}(\vec{x}, \vec{o})$$

subject to

constraints on probabilities of initial states

$$\boldsymbol{linear}(\vec{p}) = 1,$$

constraints on probabilities of transitions in every states

$$\boldsymbol{linearmulti}(A) = 1,$$

box on variables

$$p_i \geq 0,\ \ a_{ij} \geq 0,\ \ \sigma_i \geq \varepsilon,\ \ \ i = 1, \dots, n, j = 1, \dots, n.$$

***Remark*:** variables for *hmm_normal* function and constraints on probabilities are generated by PSG automatically, so user should not define these variables and constraints in Problem statement.